

Minä haluan tehdä näin. Minä ajattelen: tämä on hyvä –
Natural Semantic Metalanguage suomenkielisessä S2-kirjallisuudessa

Tuomas Hoffrén
Tampereen yliopisto
Viestintätieteiden tiedekunta
Suomen kielen tutkinto-ohjelma
Pro gradu -tutkielma
Toukokuu 2018

Tampereen yliopisto
Viestintätieteiden tiedekunta
Suomen kielen tutkinto-ohjelma

HOFFRÉN, TUOMAS: *Minä haluan tehdä näin. Minä ajattelen: tämä on hyvä* - Natural Semantic Metalanguage suomenkielisessä S2-kirjallisuudessa
Pro gradu -tutkielma, 139 sivua + liitteet [1 kpl, 6 sivua]
Toukokuu 2018

Tutkimukseni tavoite on selvittää, miten Natural Semantic Metalanguage -teorian eli NSM:n alkusanat ovat läsnä suomenkielisessä S2-kirjallisuudessa. Alkusanat ovat joukko sanoja tai muotoja, joiden oletetaan edustavan semanttisia universaaleja, eli muotoja, jotka ovat läsnä kaikissa maailman kielissä.

NSM:n käyttö S2- ja L2-opetuksessa perustuu sille olettamalle, että alkusanat ovat **1)** helppoja oppia, **2)** lähtökohtaisesti yleisiä kielessä ja **3)** tehokkaita tilanteissa, joissa oppija ei tiedä sanaa tai käsitettä, jota haluaa kohdekielessä käyttää. Tähtäimessäni on siis tutkia, opetetaanko näitä sanoja oppijoille jo varhaisessa vaiheessa.

Aineistonani toimii yksi laajalti käytetty S2-oppikirja: Hyvin menee! 1. Aineisto on nähdäkseni riittävä tiettyjen lainalaisuuksien selvittämiseen, mutta tunnustan lisätutkimuksen tarpeen niin oppikirjojen kuin muidenkin materiaalien osalta.

Tutkimuksen teoriatausta rakentuu kahdesta suuremmasta kokonaisuudesta: NSM:stä sekä Second language acquisition -tutkimuksesta, eli SLA:sta, jolla tarkoitetaan toisen kielen oppimisen tutkimusta. Tutkimuksessani siis risteää kaksi eri tutkimussuuntaa: kielitiede ja pedagogiikka.

Tutkimusmenetelmäni jakautuu kahtia: Laadullisesti pohdin sellaisia asioita, jotka tulee ottaa huomioon NSM:ää käytettäessä ja tutkittaessa. Määrällisesti teen tämän pohdinnan perusteella tarkan tilastollisen analyysin.

Tilastollisen analyysin perusteella voin sanoa, että 76 laskemaani alkusanaa jakautuvat kirjassa jokseenkin epätasaisesti. Jotkin sanat ovat erittäin yleisiä, toiset eivät esiinny kirjassa lainkaan. Kirjan alun kappaleet ovat eri alkusanojen määrän puolesta suppeita, mutta alkusanojen määrä sekä eri alkusanojen jakauma lisääntyvät huomattavasti kirjan edetessä. Kirjan eri osa-alueista eniten alkusanoja sisältävät tekstit, mutta myös tehtävissä niitä on paljon.

Tutkimukseni toimii eräänlaisena alustustyönä mahdolliselle alkusanoja hyödyntävälle opetusmetodille. Se siis rakentaa kuvan nykytilanteesta, jonka päälle uutta voi rakentaa.

Avainsanat: NSM, SLA, oppikirjat, S2- ja L2-opetus, *communication strategies*

Taulukot:

Taulukko 1 – alkusanojen suhde kirjan kokonaissanamäärään	89
Taulukko 2 – ihminen, olio ja niihin liittyvät sanat	107
Taulukko 3 – muut nominit	108
Taulukko 4 – määrä	110
Taulukko 5 – adjektiivit	112
Taulukko 6 – verbit	117
Taulukko 7 – aika	120
Taulukko 8 – paikka	122
Taulukko 9 – kieliopilliset sanat	123
Taulukko 10 – taipumattomat sanat	124
Taulukko 11 – alkusanojen top3 kappaleittain	125
Taulukko 12 – alkusanojen top3 sanoittain	126

Diagrammit:

Diagrammi 1 – alkusanojen määrä aineistossa 1/5	90
Diagrammi 2 – alkusanojen määrä aineistossa 2/5	92
Diagrammi 3 – alkusanojen määrä aineistossa 3/5	94
Diagrammi 4 – alkusanojen määrä aineistossa 4/5	95
Diagrammi 5 – alkusanojen määrä aineistossa 5/5	96
Diagrammi 6 – alkusanojen kokonaismäärä osa-alueittain	97
Diagrammi 7 – alkusanojen suhteellinen määrä osa-alueittain	98
Diagrammi 8 – osa-alueiden varianssiarvo	100
Diagrammi 9 – osa-alueiden koonti	101
Diagrammi 10 – alkusanojen kokonaismäärä kappaleittain	102
Diagrammi 11 – alkusanojen suhteellinen määrä kappaleittain	103
Diagrammi 12 – alkusanojen keskimääräinen varianssi kappaleittain	104
Diagrammi 13 – ihminen, olio ja niihin liittyvät sanat	106
Diagrammi 14 – muut nominit 1/2	107
Diagrammi 15 – muut nominit 2/2	108
Diagrammi 16 – määrä 1/2	109
Diagrammi 17 – määrä 2/2	110
Diagrammi 18 – adjektiivit 1/2	111
Diagrammi 19 – adjektiivit 2/2	112

Diagrammi 20 – verbit 1/5	113
Diagrammi 21 – verbit 2/5	114
Diagrammi 22 – verbit 3/5	115
Diagrammi 23 – verbit 4/5	116
Diagrammi 24 – verbit 5/5	117
Diagrammi 25 – aika 1/2	119
Diagrammi 26 – aika 2/2	120
Diagrammi 27 – paikka 1/2	121
Diagrammi 28 – paikka 2/2	122
Diagrammi 29 – kieliopilliset sanat	123
Diagrammi 30 – taipumattomat sanat	124
Diagrammi 31 – alkusanat määrien mukaan	128

Liitteet:

Liite 1 – NSM:n suomenkielisten sanojen kaanon, eli alkusanat tyypillisissä käyttöyhteyksissään (Vanhatalo & Tissari 2017: Virittäjän verkkojulkaisu)	140
--	-----

SISÄLLYS

1. JOHDANTO	7
1.1 Mikä on Natural Semantic Metalanguage?	8
1.2 Tutkimuksen merkitys	11
1.3 Tutkimuskysymykset ja hypoteesit sekä tutkimuksen rakenne	12
2. NATURAL SEMANTIC METALANGUAGE – NSM	15
2.1 NSM:n suomenkielinen käännös	17
2.1.1 Kääntämisen ongelmat	19
2.1.2 Alloleksisyys ja molekyylit	23
2.2 Käyttö	25
2.3 Kritiikkiä NSM:ää kohtaan	27
2.3.1 NSM:n kulmakivet	30
2.3.2 Reduktiivisen parafraasin mielekkyys	41
2.3.3 Strong Lexicalization Hypothesis ja primitiivien olemassaolo	47
3. SECOND LANGUAGE ACQUISITION	52
3.1 SLA:n näkemys universaaleihin	54
3.2 Semanttiset universaalit ja SLA-tutkimus – NMS opetuksessa	58
4. AINEISTO	61
4.1 Kirjan esittely	62
5. MENETELMÄT.....	63
5.1 Syyt ratkaisujen taustalla	64
5.1.1 Alloleksisyys ja taivutus	65
5.1.2 Homonymia ja polysemia	66
5.2 Käyttämäni alkusanalista	67
5.2.1 Ihminen, olio ja niihin liittyvät sanat	69
5.2.2 Muut nominit	70
5.2.3 Määrä	72
5.2.4 Adjektiivit	73
5.2.5 Verbit	74
5.2.6 Aika.....	79
5.2.7 Paikka	82
5.2.8 Kieliopilliset sanat	84

5.2.9 Taipumattomat sanat	85
5.2.10 Kysymyssanat – epävirallinen kategoria	87
6. TULOKSET	88
6.1 Yleiskuva NSM-sanoista kirjoissa	88
6.2 Osa-alueiden välinen vertailu	97
6.3 Kappaleiden välinen vertailu	102
6.4 Alkusanakohtainen jakauma	105
6.4.1 Ihminen, oli ja niihin liittyvät sanat	106
6.4.2 Muut nominit	107
6.4.3 Määrä	109
6.4.4 Adjektiivit	111
6.4.5 Verbit	113
6.4.6 Aika	119
6.4.7 Paikka	121
6.4.8 Kielipilliset sanat	123
6.4.9 Taipumattomat sanat	124
6.4.10 Kappaleiden top3-listat	125
6.5 Oleellisimmat tulokset	128
7. PÄÄTELMÄT JA POHDINNAT	130
8. LÄHTEET	134
9. LIITTEET	140

1. JOHDANTO

Tutkimukseni tavoitteena on tutkia Natural Semantic Metalanguage -teorian, eli NSM:n, käyttöä suomi toisena kielenä -opetuksessa. Tutkimusta tämän aiheen piiristä on tehty vähän, joten sen muotoileminen ei ole itsestään täysin selkeää. Teorian sisällä on kuitenkin esitetty, että sen käytöstä voisi olla hyötyä toisen kielen opetuksessa, ja tämä on työni ensimmäinen ponnistuslauta. Työni ei kuitenkaan suoraan rakenna mallia, miten NSM:ää voisi käyttää opetuksessa. Se on sen sijaan ensiaskel, jonka tarkoituksena on erityisesti hahmottaa nykyinen tilanne: Jotta voidaan hyödyntää jonkin teorian puitteissa tehtyä tutkimusta, on hyvä ensin tuntea, kuinka paljon teorian piirissä oleva data on jo opetuksessa läsnä.

Tarkalleen ottaen tarkastelen siis yhtä suomenkielistä paljon käytettyä oppikirjaa, ja tutkin, kuinka paljon ja missä yhteyksissä NSM:n piirissä *alkusanoina* tunnetut sanat esiintyvät tässä kirjassa. Tuotan siis tilastollisen analyysin näiden sanojen kerrostumisesta kyseiseen kirjaan.

Pelkkä tilastollinen katsaus ei kuitenkaan ole mielestäni riittävä. Jotta voidaan ymmärtää ilmiötä paremmin, on syytä mennä pintaa syvemmälle. Tilastollisen tiedon lisäksi tutkimuksellani on myös laadullinen aspekti. NSM:n alkusanat on alun perin tehty englanniksi, eikä niiden kääntäminen suomeksi ole mikään itsestäänselvyys. Pohjaan työssä käyttämäni käännökset Vanhatalon, Tissarin ja Idströmin artikkeliin *Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish* vuodelta 2014 (SKY Journal), jossa he esittelevät käännökset näille englanninkielisille sanoille. Olen kuitenkin aineistoani käsitellessäni törmännyt muutamiin kohtiin, jotka tarvitsevat vielä lisäpohtimista. Näihin huomioihin keskityn tarkemmin menetelmiä käsittelevässä luvussa 5.

Aloitan tutkimukseni pureutumalla teoriaan. Ensimmäiseksi esittelen NSM:ää teoriana, sekä sitä, mitä sen puitteissa tehdään. Esittelen luvussa yksityiskohtaisesti Vanhatalon, Tissarin ja Idströmin (2014) käännökset alkusanoille ja referoin artikkelista heidän huomionsa, jotka ovat johtaneet niihin päätöksiin, joita he ovat tehneet. NSM on saanut osansa myös kritiikistä, ja tutkimuksessani esittelen myös sitä.

Seuraavassa luvussa esittelen Second Language Acquisition eli SLA-tutkimusta ja sitä, miten NSM:n piirissä oleelliset universaalit nähdään sen puitteissa ja miten ne mahdollisesti voisivat liittyä toisen kielen opetukseen.

Tämän jälkeen esittelen aineistoni ja perustelen sen valinnan. Esittelen myös niitä ratkaisuja, joita olen aineistoni käsittelyyn liittyen tehnyt. Aineiston esittelyä puolestaan seuraa menetelmieni läpikäyminen. Kirjoitan siis tarkkaan auki ne keinot, millä olen tulokset aineistostani kerännyt. Tämä on hyvä tehdä tarkkaan, sillä aineistoa käsitellessäni olen löytänyt melko paljon asioita, jotka vaativat selkeää auki kirjoittamista. Kiinnitän menetelmäluvussa myös huomiota niihin syihin, miksi suomenkielinen NSM:n käänös vaatii erityistä huomiota. Katson, että menetelmäluku edustaa työni laadullista aspektia, sillä pohdin siinä myös alkusanojen erilaisia ominaisuuksia ja sitä, mitkä muodot edustavat suomen kielessä mitäkin alkusanaa tai semanttista primitiiviä.

Menetelmälukua seuraa luku, jossa esittelen tulokseni. Esitän sen neljässä osassa: Pyrin suuntaamaan yleisistä huomioista yksityiskohtaisempiin, eli aloitan kuvaamalla kirjan alkusana-kerrostumia yleisesti. Tämän jälkeen vertailen kirjan eri osa-alueita, sitten kappaleita ja lopuksi tarkastelen kutakin alkusanaa erikseen. Luvussa 6.5 kokoan yhteen koko tulosluvun oleellisimmat tulokset.

Lopulta viimeisessä luvussa tarkastelen kaikkia tuloksiani ja pohdin, mitä niistä on pääteltävissä. Otan huomioon syitä, jotka ovat johtaneet nykytilanteeseen ja myös sitä, miten tästä eteenpäin olisi hyvä jatkaa. Erityistä huomioita ansaitsee myös kysymys siitä, pitäisikö NSM:ää hyödyntää opetuksessa ja opetusmateriaaleja tehtäessä, ja miten tätä voisi tehdä.

Tutkimuseettisten kysymyksien pohdinta ei ole työssäni erityisen relevanttia, sillä se pohjaa kokonaan teoreettisen viitekehyksen soveltamiseen sekä kirja-aineiston tutkimiseen. Olen saanut aineistonani toimivan oppikirjan suoraan kustantajalta ja luvannut raportoida heille tutkielmani tuloksista. Olen pyrkinyt parhaani mukaan kunnioittamaan tehtyjä ratkaisuja, enkä ole käsitellyt aineistoa arvottavasti. Olen myös pyrkinyt hyvään tieteelliseen käytäntöön, jonka puolesta olen pyrkinyt tekemään selvän eron lainattujen ja omien ajatusteni välille.

1.1 Mikä on Natural Semantic Metalanguage?

Natural Semantic Metalanguage, eli lyhyemmin NSM, on Anna Wierzbickan kehittämä semanttinen lähestymistapa, jolla selitetään monimutkaisia ja ennen kaikkea kulttuurisidonnaisia käsitteitä ja ilmiöitä yksinkertaisemmin parafraaseja, kiertoilmauksia, käyttäen. NSM pohjaa ajatukseen siitä, että jokaisessa kielessä on olemassa tiettyjä kiteytyneitä merkityksellisiä elementtejä, jotka voidaan esittää kielellisin keinoin ja joita ei voida hajottaa enää pienempiin osiin. Näitä elementtejä kutsutaan

alkusanoiksi (eng. *Semantic primes*). Sanat on alun perin tehty englanniksi, mutta Vanhatalo, Tisari ja Idström (2014) esittävät niille suomenkieliset käännökset. Alkusanoja käyttäen voidaan esittää niin kielellisesti kuin kulttuurisesti monimutkaisia kokonaisuuksia yksinkertaisemmiksi puretuina. NSM on siis ensisijaisesti väline kielitieteelliseen ja kulttuurilliseen analyysiin.

Natural Semantic Metalanguage, eli lyhyemmin NSM, on teoria, jonka alkuperä on Itä-Euroopassa ja puolalaisen tutkijan Andrzej Bogusławskin töissä. Sen suurimpia eteenpäin viejiä on kuitenkin puolalainen Anna Wierzbicka, joka on jatkanut työtään 1970-luvulta Australian Natural Universityssä Canberrassa (Goddard 1998a).

NSM:n semanttisen analyysin perusajatus on hajottaa tai purkaa sanoja tai käsitteitä yksinkertaisemmiksi, helpommin ymmärrettäviksi ja käännettäviksi selityksiksi, *eksplikaatioiksi*. Tätä toimintatapaa nimitetään NSM:ssä *reduktiiviseksi parafraasiksi*. Goddardin mukaan (1998a) ainoa ehdoton edellytys eksplikaation tekemiseen on se, että se koostuu semanttisesti yksinkertaisemmista sanoista kuin alkuperäinen sana itse on. Tästä kuitenkin päästään NSM:n toiseen perusajatukseen.

NSM:n piirissä on esitetty käsite *semantic primitives*, eli *semanttiset primitiivit* tai *alkusanat*. Näillä tarkoitetaan joukkoa semanttisesti erittäin yksinkertaisia sanoja, joiden ensinnäkin oletetaan esiintyvän universaaleina kaikissa maailman kielissä ja toisekseen oletetaan olevan yksinkertaisuudessaan niin alkukantaisia, että niitä ei voi enää hajottaa pienemmiksi reduktiivisen parafrasoin keinoin.

Jatkan NSM:stä tarkemmin luvussa 2, mutta esitän vielä tutkimukseni opetuksellisen aspektin lähtökohdan: NSM:n piirissä on esitetty, että sitä voisi hyödyntää toisen kielen opetuksessa. Tästä on maininnut esimerkiksi Ulla Vanhatalo luentosarjallaan Tampereen yliopistossa syksyllä 2014. Ensinnäkin NSM:n sisältämästä alkusanastosta voisi muodostaa opiskelijalle paketin, joka on hyödyllinen, kun yritetään muotoilla sanoja tai käsitteitä, joille oppijalla ei vielä ole sanaa. Esimerkiksi: Mikä on kahvikuppi? Se on *se asia, josta ihmiset juovat kahvia*. Ajatus siis on, että oppijat osaisivat toisaalta selittää käsitteitä helposti uudella opiskelemallaan kielellä, ja toisaalta se, että he voisivat vuorovaikutustilanteessa muotoilla tarkoittamansa käsitteen siten, että muut tilanteessa olevat sen ymmärtävät. Alkusanoista on siis muodostettavissa eräänlainen yleiskieli, jolla selittää monimutkaisia ja kulttuurisidonnaisiakin ilmiöitä.

Jos NSM:ää haluaa käyttää opetuksessa, on sen hyödyllisyys tarpeen perustella. Artikkelissaan *Semantic Primes and Cultural Scripts in Language Learning and Intercultural Communication* vuodelta 2007 Anna Wierzbicka ja Cliff Goddard pohtivat NSM:n olemusta L2-opetukseen soveltuvana. He siteeraavat artikkelinsa alkuun Michael McCarthyä seuraavasti:

”The idea that there might be a core or basic vocabulary of words at the heart of any language is quite an appealing one to language educators, for if we could isolate that vocabulary we could equip learners with a survival kit of core words that could be used in virtually any situation... (McCarthy 1990: 49).”

Ajatuksena tässä sitaatissa on siis se, että jos kielistä on rikastettavissa jonkinlainen perussanasto, voimme tarjota kielenoppijoille listan sanoja, joita käyttää käytännössä kaikissa tilanteissa. Tällaisen sanapaketin hyödyllisyyttä puoltaa esimerkiksi NSM:ään vahvasti sisältyvä ajatus siitä, että monimutkaisetkin sanat ja käsitteet on hajotettavissa pienempiin – ja siten toki koottavissa takaisin. Tästä tosin lisää luvussa 2.

NSM:n avulla olisi siis luotavissa eräänlainen minisanaston ja minikielioppi, joka edustaisi kaikkien kielten risteyskohtaa (Vanhatalo 2014.) Tätä minikieltä voitaisiin käyttää käsitteellisenä *lingua francana* merkityksen tutkimisessa ja selittämisessä sekä eri kielten välillä että yhden kielen sisällä (ibid.) Tämä minisanasto olisi siitä syystä muodostettavissa, että NSM:ään kuuluvat sanat sisältyvät lapsen kieleen jo varhaisessa vaiheessa (Goddard 2001.) NSM:n alkusanojen pitäisi siis olla jo valmiina ensimmäisen kielen mentaalileksikossa (Goddard & Wierzbicka 2007: 5), mikä tarkoittaa sitä, että oppijoilla olisi jo ymmärrys niiden merkityksestä. Välittääkseen tämän tiedon kohdekielille ei siis tarvitsisi nähdä yltiöpäisesti vaivaa.

NSM:n hyödyllisyyttä oppijoiden käytössä voidaan perustella myös sillä, että kuten Goddard ja Wierzbicka (2007: 15) kirjoittavat, näennäisesti helposti käännettävä sana kuten *sister* saattaa käännettäessä saada aikaan epätietoisuutta: Joissain maailman kielissä on eri sanat vanhemmalle ja nuoremmalle siskolle, sillä tällainen piirre voi olla kulttuurisesti hyvinkin merkittävä. Käännökset eivät siis aina sovi yksi yhteen, jolloin niitä on tarpeen purkaa ja selittää sanastolla, joka sopii paremmin yksi yhteen.

Vaikka NSM:n sanaston voisi olettaa yleisyytensä vuoksi kuuluvan jo varhaisiin oppimääriin, näin ei ilmeisestikään ole (Goddard & Wierzbicka 2007: 5.) Goddard ja Wierzbicka esittelevät kolme englantia puhuville kielenoppijoille tarkoitettua kirjaa (*Deutsch Heute* [Moeller ym. 2005], *The New Practical Chinese Reader* [NOCFL 2004] ja *Korean I* [Language Education Institute, Seoul National University 2005]), joista he tutkivat 500 ensimmäistä uutta sanaa. Näihin

500:aan ensimmäiseen sanaan kuului parhaimmillaankin vain noin kaksi kolmasosaa NSM:n alkusanoista. Oma tutkimukseni asettuu tähän jatkumoon.

NSM:n kotisivujen mukaan (<https://intranet.secure.griffith.edu.au/schools-departments/natural-semantic-metalanguage>) NSM:ää on sovellettu jo yli 30 kieleen ja usealle eri tieteenalalle. Kuitenkin Ulla Vanhatalon (esim. luentosarja 2014) mukaan NSM:ää on käytetty toisen kielen opetukseen erittäin vähän. NSM:n alkusanojen ilmaantumisesta lapsen kieleen on tutkittu (esim. Goddard 2001) – NSM:n alkusanat ilmaantuvat lapsen kieleen jo varhaisessa vaiheessa, jolloin ne ovat käytettävissä ja ymmärrettävissä. Tällöin voidaan siis olettaa, että kyseiset sanat kuuluvat normaalisti oppijoiden leksikkoon joka tapauksessa, ja sisällyttämällä niiden kohdekielen mukaiset käännökset oppijoiden oppiaineeseen, voidaan suhteellisen vaivatta avata portti näiden sanojen käytölle.

1.2 Tutkimuksen merkitys

Katson tutkimukseni olevan merkittävä sekä tieteellisen tutkimuksen itsensä että sen tulosten sovellettavuuden kannalta. Tieteellisen tutkimuksen kannalta tutkimukseni merkittävyyttä perustelen ensinnäkin sillä, että tutkimusta tästä asiasta on vähän, jos ollenkaan. Näin ollen tutkimus täydentää sekä toisen kielen opetuksen että NSM:n itsensä kenttää. NSM:n osalta tutkimukseni toimii myös täydennyksenä siihen suureen määrään kieliä, joita on NSM:n piirissä tutkittu. Suomi on jo näiden kielten joukossa (ks. esim., Vanhatalo, Tissari, Idström 2014), mutta tutkimukseni tarjoaa toivottavasti lisänäkökulmia myös tähän asiaan. Lisäksi pohdin jonkin verran NSM:n alkusanojen suomenkielisiä käännöksiä NSM:ää koskevassa luvussa 2.1 ja menetelmiäni koskevassa luvussa 5. Pidän uuden tiedon tuottamista myös itseisarvona, vaikka välineellinen arvo tutkimuksella mielestäni suuri plussa onkin.

Tulosten sovellettavuutta perustelen ensinnäkin toisten ja vieraiden kielten opetuksen lisääntyvällä tarpeella. Kieliä opiskellaan monissa eri instituutioissa ja moneen eri tarpeeseen: Yliopistossa kieltä saatetaan opiskella oman mielenkiinnon vuoksi tai uramahdollisuuksien kartuttamiseksi. Kotoutumiskoulutuksessa puolestaan kielen opetus tähtää maahanmuuttajien yhteiskunnallisten oikeuksien ylläpitämiseen (Aikuisten maahanmuuttajien kotoutumiskoulutuksen opetussuunitelman perusteet, 2012.) Kun otetaan huomioon, että muita kieliä kuin suomea äidinkielenään puhuvien määrä Suomessa kasvaa koko ajan, on mielestäni S2-opetuksen kehittämiseen tähtäävä tutkimus vähintäänkin perusteltua tästäkin näkökulmasta.

Koska perustelen tutkimustani S2-opetuksen kasvavalla tarpeella, on hyvä luoda pieni katsaus S2-tutkimusalaan. Ala on vielä nuori: Ensimmäinen S2-alan väitöskirja on Maisa Martinin *The Map and the Rope. Finnish Nominal Inflection as a Learning target* vuodelta 1995. Vuosina 1995-2005 ilmestyi viisi alan väitöskirjaa, mutta tämän jälkeen tahti on kiihtynyt: vuosina 2005-2015 väitöskirjoja on ilmestynyt jo vajaa pari kymmentä. S2-alalta professuurit on jo Helsingissä ja Jyväskylässä. Yhteisiä teemoja S2-tutkimuksessa on tähän mennessä ollut oppijoiden näkökulmasta vaikeat rakenteet, vuorovaikutus, siirtovaikutus, virheanalyysi ja arviointi.

Myös S2-oppikirjoja on tutkittu vain vähän. Johanna Tannerin (2012) väitöskirjassa tutkittiin kohteliaisuutta S2-oppikirjojen dialogeissa ja Marjut Vehkasen (2015) väitöskirjassa tutkittiin vuosien 1866 ja 1953 välillä käytettyjä S2-oppikirjoja. Vehkanen tutki mm. sitä, keitä olivat näiden kirjojen tekijät ja oletetut yleisöt, mikä on kielioopin rooli kirjoissa, miten erilaisia tekstityyppejä analysoidaan kirjoissa, millaisia maailman muutoksia tekstuaalisesta kontekstista on havaittavissa sekä sitä, miten kielenopetusmetodien kehitys näkyy kirjoista.

Oppikirjojen tutkimista perustelen sillä, että vaikka eritoten S2-ala on materiaaliensa puolesta hajanainen, on oppikirjat rakennettu siten, että on tietyt puitteet joiden mukaan opetusta rakentaa. Oppikirjat muodostavat siis varman syöteaineksen opiskelijoille. On toki huomattava, että opettajat tekevät itse jatkuvasti materiaalia, eikä tätä ole aina saatavilla tutkittavaksi (esim. Tanner 2012).

Huomaan vielä, että vaikka teen tutkimukseni suomen kielen opetuksen puitteissa, en näe mitään syytä sille, etteikö tutkimustani voisi hyödyntää myös L2-alalla yleisesti. Esimerkiksi menetelmä, jossa NSM:n alkusanat sisällytetään oppijoiden oppiaineeseen jo varhaisessa vaiheessa ja jossa heitä ohjataan myös käyttämään niitä, ei ole millään tavalla suomen kieleen sidottu.

1.3 Tutkimuskysymykset ja hypoteesit sekä tutkimuksen rakenne

Kuten jo painottanutkin, NSM:n käyttöä toisen tai vieraan kielen opetuksessa on tutkittu vain vähän. Tämä on tärkeää, sillä nykyisen tilanteen selvittäminen on katsoakseni ensiarvoisen tärkeää, jos haluaa pohtia NSM:n sovellettavuutta opetuksen saralla. Tarkoitan tällä sitä, että pohdittaessa tällaista lähestymistapaa se täytyy voida pohjata jo tunnettuun tietoon. Jos siis aiotaan testata sitä, minkälaisia mahdollisuuksia NSM:n alkusanojen käytössä opetuksessa voisi olla, täytyy ensin pohtia, miten ne tulisi sisällyttää oppiaineeseen. Jotta puolestaan voidaan pohtia tätä, on hyvä tarkastella, kuinka ne ovat jo edustettuina tässä aineksessa. Tällä kysymyksellä on muutama eri aspekti:

Ensinnäkin on mielenkiintoista tarkastella, mitkä alkusanat kirjallisuudessa esiintyvät ja mitkä eivät. Tämän lisäksi on kuitenkin hyvä huomata näiden sanojen laadullinen näkökulma: Tulisiko kaikkien sanojen edes kuulua oppijoille tarjottavaan ainekseen? Mitkä ovat mahdollisesti hyödyllisiä, mitkä eivät? Minkälaisissa yhteyksissä sanat kirjallisuudessa esiintyvät? Mistä niiden jakautuminen voi johtua?

Tällä perusteella esitänkin työlleni kaksi näkökulmaa, joihin myös tutkimuskysymyksiä pohjaavat. Ensimmäinen näkökulma on luonteeltaan kvantitatiivinen, mekaaninen selvitystyö, jossa selvitän alkusanojen jakaumaa kirjallisuudessa. Toinen näkökulma on yksityiskohtaisempi ja kvalitatiivinen, johon siis pyrin esittämään vastauksia lähinnä menetelmäluvussani 5. Näkökulmat ovat siis seuraavat:

- NSM-teorian alkusanat suomenkielisessä S2-kirjallisuudessa
- NSM-teorian alkusanojen soveltuvuus toisen kielen oppimiseen

Kahden näkökulman ottamista perustelen sillä, että ne tarvitsevat toisiaan muodostaakseen toimivan kokonaisuuden: Jotta voidaan selvittää, miten NSM-teorian alkusanat ovat läsnä kielenoppijoille tarjottavassa syötteessä, on syytä myös perehtyä niihin tarkemmin ja yksityiskohtaisemmin. Toisaalta, jos halutaan pohtia NSM-alkusanojen soveltuvuutta oppikirjallisuuteen ja -ainekseen, on niitä hyvä tarkastella myös siinä ympäristössä, jossa ne luonnollisesti esiintyvät. On syytä ottaa huomioon, että oppikirjallisuus on tiettyä tarkoitusta varten luotua tekstiä, ei siis välttämättä täysin luonnollista tai sujuvaa. Tästä syystä käytän lähteenäni nimenomaan oppikirjatekstiä, enkä esimerkiksi uutistekstiä. Näihin asioihin vedoten esitän tutkimuskysymykseni seuraavanlaisiksi:

1. Mitkä NSM-alkusanat esiintyvät valitussa S2-oppikirjassa ja kuinka yleisiä ne ovat?
2. Millaisissa yhteyksissä esiintyvät sanat ovat ja miten niitä käytetään? Miksi?

Ensimmäiseen tutkimuskysymykseen minulla on kaksi ennakko-oletusta: Ensinnäkin NSM:n alkusanat ovat luonteeltaan yksikertaisia, ja kuten Goddard (2001) on todennut, monet niistä sisältyvät oppimaamme kieliainekseen jo varhain. Tästä syystä voisin olettaa niiden esiintyvän myös S2-oppikirjallisuudessa jokseenkin taajaan. Toisaalta, kuten jo johdannossa mainitsin, muiden muita kieliä koskevassa L2-tutkimuksessa on havaittu, että osa alkusanoista ei sisälly varhaisiin oppikirjoihin lainkaan.

Vastauksen toiseen kysymykseen oletan olevan monimutkaisempi. NSM:n alkusanaalista sisältää mm. sellaisia sanoja kuin *alla* ja *päällä*. Kokemukseni mukaan tällaisten postpositioiden opetus tapahtuu oppikirjoissa usein ryppäissä, eli samaan aikaan esitellään useita vastaavia sanoja. Toisaalta NSM:n alkusanojen EI ja OLLA oletan esiintyvän oppikirjoissa jotakuinkin koko ajan. Lisäksi esimerkiksi alkusanaan KUOLLA en usko törmääväni oppikirjoissa lainkaan. Tämän tutkimuskysymyksen käsittelyn ohessa on mielestäni oleellista tarkastella myös syitä siihen, miksi alkusanat esiintyvät niin kuin esiintyvät.

Lisäksi haluan esittää erikseen kaksi kysymystä, joihin pyrin tutkielmassani pohtimaan vastauksia:

- Mitkä alkusanat voisivat olla hyödyllisiä toisen kielen opetusta ajatellen ja mitkä eivät?
- Millaisia erityistä huomiota vaativia piirteitä suomen kielessä on NSM-alkusanoja ajatellen?

Ensimmäisen kysymyksen vastauksen oletan olevan luonteeltaan tulkinnanvarainen ja paljolti ihmisestä riippuva. On mielestäni selvää, että kaikki alkusanat eivät voi olla opetuksessa yhtä hyödyllisiä kuin toiset. Ei esimerkiksi vaikuta relevantilta opettaa sanaa, joka ei esiinny luonnollisessa suomen kielessä juuri koskaan. Toisaalta jotkut sanat kuuluvat aineistooni, ja mahdollisesti myös luontaiseen kielenkäyttöön niin taajaan, että niihin tuskin on tarpeen kiinnittää tehostettua huomiota opetusmateriaaleja valmistettaessa. Tätä kolmatta tutkimuskysymystä aion pohtia aineistoni perusteella, sekä lisäksi aion verrata aineistostani saamiani tuloksia suomen kielen taajuussanastoon (esim. Saukkonen ym. 1979).

Jälkimmäinen kysymys pohjaa siihen, että englantia, jonka perusteella NSM-lista on tehty, on lähtökohtaisesti erilainen kieli kuin suomi. Kielillä on huomattavia eroja niin sanastonsa kuin kieliopillisten rakenteidensa perusteella. On siis syytä olettaa, että NSM-alkusanojen käännoissä on noussut esiin asioita, jotka vaativat vielä lisäpohdintaa. Tätä pohdintaa ovatkin jo Vanhatalo, Tissari ja Idström tehneet artikkelissaan vuodelta 2014, ja aion itse lisätä tähän pohdintaan sellaisia asioita, joita aineistoa käsitellessäni huomaan.

2. NATURAL SEMANTIC METALANGUAGE – NSM

Natural Semantic Metalanguage eli NSM on teoria, joka sisältää ajatuksen universaaleista semanttisista primitiiveistä ja siitä, että näitä primitiivejä käyttämällä voidaan muodostaa selitys, *eksplikaatio*, mistä tahansa sanasta tai käsitteestä. Tällaista metodologia, kuten NSM:n nimestäkin käy ilmi, voisi kuvailla siis vaikkapa metakieleksi.

Käytännössä tämä metakieli siis koostuu määrittelystä riippuen noin 65:stä alkusanasta (*primes*), joille on ominaista se, että niitä ei voi selittää tai määritellä muin sanoin: Ne ovat atomimaisia kuin matematiikan alkuluvut. NSM:n mukaan nämä alkusanat ovat löydettävissä kaikista maailman luonnollisista kielistä.

Toinen NSM:n perusajatus on jo johdannossakin mainittu reduktiivinen parafraasi. Se on metodi, jolla sana tai käsite pyritään esittämään mahdollisimman yksinkertaisena (joskin usein pitkänä) eksplikaationa. Metodi pohjaa ajatukseen siitä, että mikä tahansa termi on selitettävissä siten, että sen jakaa pienempiin, helpommin ymmärrettäviin yksiköihin.

Yksi reduktiivisen parafraasin tärkeimpiä tehtäviä on välttää *sirkulaatiota* eli kehäpäätelmiä. Tällä tarkoitetaan kehämäistä määrittelyä, jossa yksi sana määritellään käyttämällä toista, ja tämä toinen sana määritellään käyttämällä ensimmäistä. Esimerkiksi Vanhatalo, Tissari ja Idström esittävät artikkelissaan *Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish* (2014) että *The Oxford English Dictionary* määrittelee sanan *will* osittain käyttämällä sanaa *desire*.

Alkusanoja käytetään siis muodostamaan monimutkaisemmista ilmiöistä yksinkertaisempia selityksiä. Esimerkiksi angloamerikkalaista ajatusta henkilökohtaisesta itsenäisyydestä on kuvattu alkusanoilla seuraavasti:

”Monet ihmiset ajattelevat näin:

Kun joku tekee jotain, on hyvä, jos hän voi ajatella näin:

”Teen näin, koska tahdon tehdä näin.””

(Goddard 2011: 397)

Eksplikaatit muistuttavat siis tavanomaista, joskin yksinkertaista kieltä. Ideaaleina niiden tulisi olla niin osuvia, että eksplikaation voisi missä tahansa tilanteessa sijoittaa siitä tehdyn sanan paikalle ilman minkäänlaista poikkeusta merkityksessä. Tätä Goddard (1998a) kuvaa termillä *substitutability*; korvattavuus.

Kuten johdannossa sanoin, NSM:n ajastus sellaisenaan tulee puolalaiselta lingvistiltä Andrzej Bogusławskilta. Alun perin NSM-teorian ja käsitteen *alkusana* esitteli Anna Wierzbicka kirjassaan *Semantic Primitives* vuodelta 1972. Periaatteessa kuitenkin ajatus kielen pienistä yksiköistä löytyy jo Platonilta: ”Voi hyvä ystävä, etkö tiedä, että alkuperäiset sanat ovat hautautuneet kaiken alle, kun niistä on haluttu tehdä näyttävämpiä?” (Kratylos, 256). Gottfried Wilhelm Leibniz (1646-1726) puolestaan esitteli ajatuksen ajattelun yksiköistä, mistä on kirjoittanut Bertrand Russell teoksessaan *History of Western Philosophy* (1945). Eugene Nida ja Charles Taber puolestaan esittelivät (1969, 1982) ajatuksen kielen *kerneleistä*, ytimistä, jotka ovat hautautuneet monimutkaisempiin käsitteisiin.

Teoksessaan *Semantic Theory and Semantic Universals* (1994) Cliff Goddard esittää NSM:n semanttisena peruskivenä toimivat periaatteet. Hän määrittelee nämä periaatteet seitsemään erikseen auki kirjoitettavaan sääntöön.

1) *Semiotic Principle*. ”A sign cannot be reduced to or analyzed into any combination of things which are not themselves signs; consequently, it is impossible to reduce meaning to any combination of things which are not themselves meaning. (C. S. Peirce 1932: 2, 230-231)”

Tässä periaatteessa on nähtävissä se kielessä oleva perusajatus, että lopulta kieli on aina vain merkki. Se ei siis koskaan todella vastaa puhdasta ajatusta. Käytetty kieli on aina *langue*, ei koskaan *parole*. On hyvä jo tässä vaiheessa huomata, että NSM:n piirissä todellakin tiedostetaan tämä tosiasia, sillä se on yksi taajaan kritisoitu ominaisuus NSM-teorian piiristä. Tästä lisää luvussa 2.3.

2) *Principle of Discrete and Exhaustive Analysis*. ”Meaning can be analyzed in a fully determinate way; that is, any complex meaning can be decomposed into a combination of discrete other meanings, without circulation and without residue.”

Tämä periaate ottaa selvän kannan niin merkityksen todelliseen olemukseen kuin semantiikan ja pragmatiikan eroonkin. Aiheesta lisää NSM:n osakseen saamaa kritiikkiä käsittelevässä luvussa 2.3.

3) *Semantic Primitives Principle*. ”There exists a finite set of undecomposable meaning – semantic primitives. Semantic primitives have an elementary syntax whereby they combine to form ‘simple propositions’.”

Tästä periaatteesta johtuu primitiivien olemassaolo. Jotta voisi olla mitään pysyvää merkitystä, on oltava olemassa elementtejä, joiden merkitys ei vaihtelee; eräänlaisia ankkureita.

4) *Natural Language Principle*. “Semantic primitives and their elementary syntax exist as a minimal subset of ordinary natural language.”

Normaali kieli on siis aina redusoitavissa pienempiin ja yksinkertaisempiin osiin siten, että sen merkitys pysyy.

5) *Expressive equivalence of NSMs*. “The NSMs derived from various languages will be semantically equivalent, that is, have the same expressive power. Any simple proposition expressible in an NSM based on L_1 will be expressible in an NSM based on L_2 , L_3 and so on.”

Tämä käsittää siis kielten universaalit ominaisuudet. Mikä tahansa ilmaisu on ilmaistavissa myös toisella kielellä erityisesti NSM:ää käyttäen.

6) *Isomorphism of NSMs*. “The simple propositions which can be expressed through the NSMs based on different languages will be fundamentally isomorphic.”

Ilmaisu, joka siis esitetään NSM:ää käyttäen, on täsmälleen yhteneväinen toisen kielen NSM-käännöksellä esitetyn ilmauksen kanssa.

7) *Strong Lexicalization Hypothesis*. “Every semantically primitive meaning can be expressed through a distinct word, morpheme or fixed phrase in every language.”

Jokainen semanttinen primitiivi on siis ilmaistavissa jokaisella maailman kielellä, ja ne ovat myös erityisen eksplisiittisinä näissä kielissä. Tämä oletus on saanut osakseen kritiikkiä, ja siitä lisää luvussa 2.3.3.

2.1 NSM:n suomenkielinen käännös

Ihanteellinen versio NSM:stä on täysin aukottomasti käännettävissä oleva sanalista, joka ei sisällä minkäänlaisia merkityseroja kielestä riippumatta. Todellisuudessa kuitenkin alun perin englanniksi kirjoitetun NSM:n alkusanalistan kääntäminen on vaativa tavoite. Myös tässä realisoituu se, että kielet eivät osu täysin yksi yhteen, vaan yksinkertaisiltakin näyttävillä sanoilla voi olla piilomerki-

tyksiä. Vanhatalo, Tissari ja Idström ovat esittäneet SKY Journalissa (27: 2014) julkaistussa artikkelissaan suomenkielisen käännöksen NSM:n alkusanalistaan. Lisäksi he pohtivat päänvaivaa tuottavia rakenteita sekä merkityksiä tiettyjen sanojen kääntämisen kohdalla.

Vanhatalon, Tissarin ja Idströmin kääntämä alkusanalista on seuraavanlainen:

”English (NSM homepage 2013)

Finnish (Vanhatalo & Tissari tulossa, [huom.] ks. Vanhatalo & Tissari 2017)

I	MINÄ
YOU	SINÄ
SOMEONE	JOKU~IHMINEN~HÄN
PEOPLE	IHMISET
SOMETHING~THING	JOKIN~ASIA
BODY	RUUMIS~KEHO
KIND	-LAINEN~-LÄINEN
PART	OSA
THIS	TÄMÄ~SE
THE SAME	SAMA
OTHER~ELSE	TOINEN~MUU
ONE	YKSI~ERÄS
TWO	KAKSI
SOME	JOKIN~JOKU~MUUTAMA
ALL	KAIKKI
MUCH~MANY	PALJON~MONI
LITTLE~FEW	VÄHÄN~HARVA
GOOD	HYVÄ
BAD	PAHA
BIG	ISO~SUURI
SMALL	PIENI
THINK	AJATELLA
KNOW	TIETÄÄ
WANT	TAHTOA
FEEL	TUNTEA~TUNTUA
SEE	NÄHDÄ
HEAR	KUULLA
SAY	SANOA
WORDS	SANAT
TRUE	TOSI~TOTTA
DO	TEHDÄ
HAPPEN	TAPAHTUA
MOVE	LIKKUA
TOUCH	KOSKEA
BE (SOMEWHERE)	OLLA (JOSSAIN)
THERE IS	OLLA (OLEMASSA)
HAVE ¹⁰	OLLA (OMISTAA)
BE (SOMEONE/SOMETHING)	OLLA (JOKU/JOKIN)
LIVE	ELÄÄ
DIE	KUOLLA

WHEN~TIME	MILLOIN~JOLLOIN~SILLOIN~ AIKA
NOW	NYT
BEFORE	ENNEN~AIKAA SITTEN~AIKAISEMMIN
AFTER	JÄLKEEN~AJAN KULUTTUA
A LONG TIME	KAUAN (AIKAA)~PITKÄN AJAN
A SHORT TIME	VÄHÄN AIKAA
FOR SOME TIME	JONKIN AIKAA
MOMENT	HETKI
WHERE~PLACE	MISSÄ~JOSSA~PAIKKA
HERE	TÄSSÄ~TÄÄLLÄ
ABOVE	PÄÄLLÄ
BELOW	ALLA
FAR	KAUKANA
NEAR	LÄHELLÄ
SIDE	PUOLI
INSIDE	SISÄ-
NOT	EI
MAYBE	EHKÄ
CAN	VOIDA
BECAUSE	KOSKA~VUOKSI~TAKIA
IF	JOS
VERY	HYVIN~ERITTÄIN
MORE	ENEMMÄN~LISÄÄ (and ENÄÄ)
LIKE~AS~WAY	NÄIN~KUTEN”

2.1.1 Kääntämisen ongelmat

Oma vaivansa NSM:n alkusanojen mukauttamiseen jokaiselle kielelle on siinä, että niitä määritellen usein englannin kautta. Alkuperäiset englanninkielisetkin alkusanat saattavat kantaa mukanaan erilaisia merkityksiä esimerkiksi polysemian ja käyttöympäristön kautta, joten on uskottavaa olettaa, että käännökset ja muista kielistä hahmotellut alkusanalistat sisältävät näitä myös.

Alloleksillä tarkoitetaan sanaa, jolla on sama semanttinen merkitys kuin alkusanalla, mutta joka on syystä tai toisesta erilainen. Joskus tämä tarkoittaa synonyymejä, joskus kyse on esimerkiksi sijataivutuksesta. Esimerkkeinä tilanteista, jolloin alloleksisyyttä tulisi pohtia, Vanhatalo, Tissari ja Idström (2014) esittävät alkusanat ISO-SUURI sekä HALUTA-TAHTOA. ISO ja SUURI voivat merkitä suurin piirtein samaa asiaa, tai niillä voi olla merkittävä semanttinen ero: Lauseiden *Hän on iso mies* ja *Hän on suuri mies* merkitysero on selkeä; ensimmäisellä tarkoitetaan miehen fyysistä kokoa, jälkimmäinen voi tarkoittaa myös miehen merkitystä tai hänen luonnettaan (Vanhatalo, Tissari & Idström 2014).

Englanninkielisen alkusanan WANT käännöksenä Vanhatalo, Tissari ja Idström (ibid.) käyttävät sanaa TAHTOA. Vaikka sanojen *haluta* ja *tahtoa* välillä merkitysero on joskus

pieni, em. katsovat sanan *tahtoa* neutraalimmaksi kuin sanan *haluta*. Jälkimmäisellä he katsovat olevan enemmän primitiivinen ja jopa seksuaalisesti latautunut kaiku. He esittävät määritelmänsä molemmille sanoille seuraavasti: *Tahtoa*: 'determined and purposeful action when trying to reach some goal' ja *haluta*: 'someone's action while trying to fill some need or to make some wish true'.

Englanninkielisten alkusanojen kääntäminen suomen kieleen on ongelmallista sanaston paikoin hankalan yhteensovittamisen lisäksi myös suomen kielen synteettisen luonteen vuoksi (ibid.). Hyvän esimerkin Vanhatalo, Tissari ja Idström (ibid.) tarjoavat pohtiessaan englanninkielisten alkusanojen KIND ja LIKE sovittamista suomen kieleen. Tässä tilanteessa on syytä ottaa huomioon myös homonymia: LIKE-alkusana tarkoittaa tässä yhteydessä samankaltaisuutta toiseen asiaan, ei *like*-verbiä. Homonymian lisäksi ongelmia voi tuottaa sanan polysemia. Esimerkkinä Vanhatalo, Tissari ja Idström esittävät alkusanan TEHDÄ. Tämän sanan merkitykseksi sallitaan tässä NSMä:n käännöksessä vain englannin *to do* -kaltainen merkitys: *to make* on poissuljettu.

Vanhatalo, Tissari ja Idström (ibid.) huomauttavat, että huolimatta semanttisesta samankaltaisuudesta, KIND ja LIKE ovat eri alkusanoja, ja niille on myös eri käännös. KIND alkusanalle he esittävätkin käännökseksi johdinta *-lAinen*. Jokainen englanninkielinen alkusana ei siis ole heidän mukaansa käännettävissä toiseksi lekseemiksi, vaan suomen kielen ominaisluonteen vuoksi on otettava huomioon myös johtimet ja taivutusmuodot. Sanalle LIKE he puolestaan esittävät allolekeiksi sanoja AS ja WAY, ja käännös olisi heidän mukaansa NÄIN tai KUTEN.

Käännöksellisiä ongelmia tuottavat myös englanninkieliset alkusanat ONE, SOMETHING-THING, LIVE, MORE, BAD, WANT, BE, KNOW sekä THIS ja HERE. Myös näitä ongelmia Vanhatalo, Tissari ja Idström pohtivat artikkelissaan.

Alkusanan ONE ongelmallisuuden Vanhatalo, Tissari ja Idström (ibid.) tuovat esiin englanninkielisen sanan *cup* NSM-määritelmän avulla: *Someone can hold_[m] one in one hand_[m]*. Tätä ei voi kääntää sujuvasti suomeksi käyttämällä alkusanaa ONE, sillä kuten Vanhatalo, Tissari ja Idström osoittavat käännöksessään *Ihminen voi pitää_[m] sellaista yhdellä kädellä_[m]*, ONE-alkusana saa käännökseen sanan *sellaista*, jonka nominatiivimuotoista johdinta *-lAinen* em. ehdottavat käännökseksi alkusanalle KIND.

Myös englanninkielinen *the other one* on käännetty *tuo toinen*, tai *tuo toinen tällainen*, mikä osoittaa sen, että alkusana ONE sisältää englannissa käyttötarkoituksia, jotka eivät esimerkiksi suomen kielessä enää toimikaan. Em. toteavat, että alkusanan YKSI alloleksinä toimisi myös sanan ERÄS.

Alkusanojen SOMETHING-THING ongelman Vanhatalo, Tissari ja Idström tiivistävät koskemaan suomen kielen sanoja *asia* ja *esine*: *Asialla* on totutusti tarkoitettu jonkinlaista abst-

raktia käsitettä, ja *esineen* on katsottu koskevan konkreettista, mutta suhteellisen pienikokoista objektia. Kumpikaan puolestaan ei kelpaa tarkoittamaan eläviä olentoja tai erityisen suurikokoisia objekteja, kuten kissaa tai taloa. Tällöin em. ehdottavat vaihtoehtoiseksi sanaksi alkusanaa JOKIN.

LIVE aiheuttaa Vanhatalon, Tissarin ja Idströmin mukaan päänvaivaa sen saamien käännösten *elää* ja *asua* vuoksi. He huomauttavat myös Wierzbickan todenneen, että tällaista eroa ei englanninkielisessä alkusanassa LIVE tehdä.

Alkusanojen ONE, SOMETHING-THING ja LIVE myötä on siis nähtävissä, miten polysemia voi olla esimerkiksi englanninkielisessä sanassa olemassa, mutta saattaa jäädä puuttumaan käännetyistä, jolloin on joko tarpeen hajauttaa kyseisen englanninkielisen alkusanan merkitys kahteen uuteen alkusanaan tai tarkentaa kyseisen alkusanan merkitystä, kuten on tehty alkusanassa TEHDÄ.

Alkusanan MORE merkityksistä Vanhatalo, Tissari ja Idström (ibid.) esittävät, että vaikka englanninkielisen jaon sanojen *more* ja *anymore* välillä on semanttinen ero, tämä eri ei toteudu samanlaisena näiden sanojen suomenkielisten käännösten kanssa. Kun sanan *more* voidaan katsoa tarkoittavan jonkin asian määrää, tarkoittaa *anymore* selkeästi aikaa. Puolestaan MORE:n suomenkielisten käännösten ENEMMÄN, LISÄÄ ja ENÄÄ välillä yhtä tarkkarajaista eroa ei ole (ibid.).

Englanninkielinen alkusana BAD tuottaa ongelmia niin suomessa kuin muissakin kielissä: Vanhatalon, Tissarin ja Idströmin (2014) mukaan venäjän kielessä BAD saa ainakin neljä vastinetta ja useamman myös mandariinikiinassa. Suomen kielessä em. ehdottavat vastineiksi sanoja PAHA ja HUONO sekä joskus myös TUHMA. Usein näiden sanojen semanttinen vastakohta olisi HYVÄ, joten havaittavissa on tietynlainen epätasapaino sanojen välillä.

Hyvän esimerkin englanninkielisen ja suomenkielisen NSM:n yhteensovittamisesta Vanhatalo, Tissari ja Idström (2014) osoittavat olevan BE, joka kääntyisi täten suomenkieliseksi alkusanaksi OLLA. Tämän alkusanan em. hajottavat neljään eriävään merkitykseen: ensinnäkin he määrittävät omaksi alkusanakseen paikkaa kuvaavaan OLLA (JOSSAIN). Toisekseen englanninkielistä ilmaisua *there is*, jolla kuvataan jonkin olemassaoloa em. ehdottavat kuvaamaan alkusanan OLLA (OLEMASSA). Kolmas merkitys on suomen kielen omistusta kuvaava rakenne *-lla on*, jolle em. ehdottavat vastinetta OLLA (OMISTAA). Viimeinen merkitys kuvaamaan englannin BE-alkusanaa on OLLA (JOKU/JOKIN), joka kuvaa kyseisen asian tai olion luonnetta.

Suomen kielessä on selkeä ero sillä, *tietääkö* joku toisen ihmisen vai *tunteeko* hän hänet. Vaikka molemmat sanat voivat kääntyä englanniksi sanaksi *know*, Vanhatalo, Tissari ja Idström esittävät alkusanan KNOW käännökseksi vain sanaa TIETÄÄ. Tällöin TUNTEA on sanan FEEL käännöksenä yhdessä alkusanan TUNTUA kanssa.

Englanninkielisessä NSM:ssä on erotettu toisistaan selkeästi sanat THIS ja HERE. Tämä, kuten Vanhatalo, Tissari ja Idström esittävät, ei ole suomen kielessä niin yksinkertaista, koska HERE-alkusanaa vastaavat käännökset TÄSSÄ on puolestaan THIS-alkusanaa vastaavan TÄMÄ-alkusanan inessiivimuoto. Näin ollen alkusanoja TÄSSÄ-TÄÄLLÄ ei voida määrittää ilman alkusanaa TÄMÄ, mikä puolestaan ei pidä samalla tavalla paikkaansa englanninkielisillä alkusanoilla THIS ja HERE. Huomionarvoista on myös se, että TÄSSÄ ja TÄÄLLÄ eroavat usein toisistaan myös semanttisesti, kuten ulko- ja sisäpaikallissijat suomen kielessä yleensäkin. Vanhatalo, Tissari ja Idström peräänkuuluttavatkin jatkotutkimusta siitä, onko HERE lopulta vain alkusanan THIS alloleksi.

Aikaa kuvaavien NSM-alkusanojen osalta Vanhatalo, Tissari ja Idström pohtivat esimerkiksi sitä, missä määrin ajan pituutta kuvaavat *for some time* ja *for long time* ovat täysin määriteltävissä muilla alkusanoilla. Jos näin olisi, ne eivät olisi alkusanoja. Huomionarvoista heidän mukaansa on kuitenkin se, että useat kielet, suomi mukaan lukien, sisältävät ajan ilmauksia, joissa ei käytetä sanaa *aika* lainkaan. Tällainen ilmaus on esimerkiksi KAUAN. Toisaalta vastaavaa lyhyttä aikaa kuvaavaa sanaa ei suomen kielessä ole. Myös ajan pituutta indikoiva *for* englanninkielisissä ilmauksissa aiheuttaa päänvaivaa; se sisältyy englanninkieliseen alkusanalistaan ilmauksessa FOR SOME TIME, mutta ei ilmauksissa A LONG TIME ja A SHORT TIME. Omia huomioitani asiasta luvussa 5.

Vanhatalon, Tissarin ja Idströmin (2014) mukaan alkusana SOME onkin tuottanut ongelmia usein. Tällä hetkellä sitä ei käytetä kuvaamaan juurikaan ainesanoja, sillä tällöin alkusanan SOME merkitys voitaisiin muotoilla myös *not much, not little*. SOME esiintyy kuitenkin ilmauksessa FOR SOME TIME. Edellä mainitut pohtivatkin, josko alkusanan SOME merkitykseen voisi sisällyttää myös ainesanojen määrän. Tällä hetkellä kuitenkin alkusanan SOME suomenkielisenä käännöksenä toimii JOKIN-JOKU-MUUTAMA.

Englanninkielisessä NSM-listassa esiintyy kaksi sanaa, jotka ovat jo valmiiksi monikkomuodossa: PEOPLE ja WORDS. Erityisen paljon keskustelua on aiheuttanut alkusana PEOPLE, sillä sen englanninkielinen yksikkömuoto *person* tai *someone*.

Sille, että WORDS on monikossa, Vanhatalo, Tissari ja Idström (ibid.) esittävät Goddardilta lähtöisin olevan perustelun siitä, miten WORDS esiintyy yleensä englanninkielisessä ympäristössä monikkomuodossa. He kuitenkin ehdottavat, että tämän alkusanan voisi hyvin muuttaa yksikkömuotoiseksi. Ylipäätään sanat SANAT ja IHMISET vaikuttavat suomenkielisessä kontekstissa erikoisilta, ja tämän allekirjoittaa myös Vanhatalo (sähköpostikeskustelu). Esittelen lisää alkusanoista tekemiäni huomioita menetelmäluvussani 5.

2.1.2 Alloleksisyys ja molekyylit

Alloleksi on alun perin Anna Wierzbickalta oleva termi. Tällä kuvataan tilannetta, jossa kaksi tai enemmän sanoja tietyssä kielessä vastaavat yhtä semanttista primitiiviä. Artikkelissaan *The Natural Semantic Metalanguage Approach: An overview with reference to the most important Romance languages* (2006) Cliff Goddard ja Bert Peeters esittävät esimerkiksi englannin kielen sanan *I* ja semanttisen primitiivin, joka vastaa tätä sanaa. Kuten tiedämme, tilanteesta riippuen englannin kielessä tulee samasta asiasta käyttää joko sanaa *I* tai *me*. Nämä ovat siis allolekseja; ne vastaavat samaa primitiiviä, mutta ovat muodollisesti erilaisia. Kyseessä on siis eräänlainen synonymia.

NSM:ssä tehdään ero kolmen erilaisen alloleksisyyden välillä (ibid.) Nämä tyypit ovat *positional*, *combinatorial* ja *inflectional*, eli vapaasti käännettynä *paikallinen*, *yhdistelmällinen* ja *taivutuksellinen*.

Annettu englanninkielinen esimerkki *I* ja *me* vastaavat puhtaasti ensimmäistä tyyppiä. Tässä tyypissä eri muoto siis selittyy puhtaasti sen paikan tai syntagmaattisen roolin perusteella. Toinen, *combinatorial*, puolestaan tarkoittaa sellaista alloleksisyyttä, jossa sanan suhde toisen samaan muokkaa sen muotoa, kuten esimerkiksi ranskan kielen adjektiivissa *pieni*: *petit* [masc.sg], *petite* [fem.sg], *petits* [masc.pl] ja *petites* [fem.pl]. Erityisesti tämä ilmiö siis näkyy esimerkiksi attribuuteissa sellaisissa kielissä, joissa sanoilla on suku. Suomen kannalta tämä ei ole järin oleellista.

Sen sijaan suomen kielen kannalta on erittäin oleellinen kolmas alloleksisyyden tyyppi, *inflectional*. Se tarkoittaa sanan muuttumista taivutuksen perusteella, mikä on suomen kielessä itsestään selvää. Peeters ja Goddard (ibid.) antavat siitä esimerkiksi englannin kielen tempusvaihtelun: Kuten englantia taitavat tietävät, verbi *do*, jolla on siis myös semanttinen primitiivi vastineenaan, saa menneessä muodossa muodon *did* tai *done*. Em. kuitenkin kirjoittavat, että eksplikaatiota kirjoitettaessa joku näistä muodoista voi olla yksinkertaisesti epäkieliopillinen, mikä ei toki ole eksplikaation tarkoitus. Peeters ja Goddard siis sallivat ainakin tempustaivutuksen verbien osalta. Huomionarvoista asiassa vain on sen, että suomi synteettisenä kielenä sisältää huomattavan paljon enemmän taivutusta kuin vaikkapa englanti, joka on lähes analyttinen kieli (ks. esim. Itkonen 2008, Pajunen: luentosarja 2018).

Vanhatalo, Tissari ja Idström (2014) painottavat termiä *kieliopillinen alloleksi* erotuksena allolekseihin, jotka ovat eri leksemeitä, mutta palvelevat NSM-alkusanoissa hyvin pitkälti samaa tar-

koitusta, eli ns. ensimmäisen tyyppin allolekseihin. Suomen kielessä on viisitoista sijamuotoa ja lisäksi verbeillä oma taivutussysteeminsä, joten tilanne on sikäli erilainen kuin englannissa: Mitkä sanojen taivutusmuodot tulee hyväksyä NSM-alkusanojen paradigmaan ja miksi?

Paikallissijojen osalta asia on Vanhatalon, Tissarin ja Idströmin (ibid.) mukaan sikäli ongelmallinen, että esimerkiksi elatiivin ja ablatiivin merkitys voidaan englanniksi kääntää sanaan *from* sekä illatiivin ja allatiivin merkitys voidaan kääntää sanoihin *to* ja *for*. *From*, *to* ja *for* puolestaan ovat ilmaistavissa parafraasein NSM-sanastoa käyttäen, joten ei ole automaattisesti selvää, voidaanko elatiivi-, ablatiivi-, illatiivi- ja allatiivi-muotoiset sanat hyväksyä NSM-alkusanoiksi.

MA-infinitiivin abessiivirakennetta, erityisesti ilmaisua *ei voi olla tekemättä* Vanhatalo, Tissari ja Idström ovat käsitelleet esimerkiksi siitä lähtökohdasta, että vaikka tämänkaltainen ilmaisu ei englannin kielessä ole luonteva, ei sen englanninkielinen käännös *can not not do* toimi puolestaan suomeksi. NSM-alkusanojen ketjut muodostavat joskus monimutkaisiakin kokonaisuuksia yksinkertaisista käsitteistä, mutta *ei voi ei tehdä** riitelee jokseenkin vahvasti suomen kielen kielipin kanssa. Vanhatalo, Tissari ja Idström myös huomauttavat, että vaikka voisi ajatella kyseisen ilmauksen merkitsevän myös ilmausta *täytyy tehdä*, nämä ilmaukset eivät ole semanttisesti yhtenevät. Näiden huomioiden valossa em. ehdottavat, että ilmaisu *joku ei voi olla tekemättä jotain* käännetään muotoon *someone can't be without doing something*.

Vanhatalo, Tissari ja Idström (ibid.) lopettavat pohdintansa kielipillisisistä allolekseista siihen huomioon, että NSM-paradigma saattaa tietyissä tilanteissa suosia tietynlaista kielipillistä monimutkaisuutta, ja samalla se saattaa syrjiä toista. Aihe on kiinnostava, ja em. huomauttavatkin, että asiaa tulisi pohtia vielä pidemmälle. On syytä jo tässä vaiheessa mainita, että minun tulee työni puitteissa tehdä selkeä linjaus tähän asiaan. Lähestyn aiheen osalta tekemiäni ratkaisuja tarkemmin menetelmiäni käsittelevässä luvussa 5.

Viimeinen termi, jonka tässä määrittelen, on *molekyyli*. Molekyyllillä tarkoitetaan NSM:n yhteydessä sanaa, joka ei kuulu alkusanojen listaan, mutta jolle on tehty alkusanoja käyttäen kattava määrittely. Näin ollen tilan ja vaivojen säästämiseksi sanaa käytetään sellaisenaan, mutta sen perään merkitään alaindeksiin hakasulkeisiin molekyyliä tarkoittava kirjain *m*; ”_[m]”.

Molekyyli auttavat tekemään eksplikaatioista paremmin ymmärrettäviä. Lisäksi molekyylien käyttö säätelee myös niiden pituutta. Kuten Goddard teoksessaan *Semantic Analysis: A practical introduction* (1998) sanoo, ei eksplikaatiota aina tarvitse esittää pelkästään alkusanoja käyttämällä: Oleellista on se, että siinä käytetyt sanat ovat yksinkertaisempia kuin määriteltävänä oleva sana. Huomautettakoon, että myös Piotr Blumczynski (2013) ja Nick Riemer (2006), joiden

ajatuksista kirjoitan tarkemmin NSM:n osakseen saamaa kritiikkiä käsittelevässä luvussa 2.3, pe-
räänkuuluttavat sitä, olisiko viisaampaa luoda eksplikaatiot siten, että ne edistäisivät ymmärrystä
absoluuttisen yksinkertaisuuden sijaan. Tähän kysymykseen Ulla Vanhatalon (sähköpostikeskus-
telu) mielestä molekyylit vastaavat hyvin, ja olen itsekin taipuvainen näkemään asian näin.

2.2 Käyttö

NSM:n piirissä ollaan oltu erityisen kiinnostuneita eri kulttuureihin kuuluvista käsitteistä ja sa-
noista. Tämä liittyy useampaankin pohja-ajatuksen, mutta on kiintoisaa myös oman tutkimukseni
kannalta: On kieltämättä selvää, että kulttuurisidonnaisten käsitteiden ymmärtäminen voi olla kie-
lenoppijalle hankalaa. Se liittyy niin uuden kieliympäristön mukanaan tuomiin vaikeuksiin kuin so-
pivien sanakirjakäännösten puuttumiseenkin. Rinta rinnan kulttuurisidonnaisten käsitteiden kanssa
on siis läsnä myös sanojen kääntäminen. Tämä molemmat puolestaan linkittyvät omaan työhöni
opetuksen tiimoilta.

Wierzbicka esittää fraasinomaisesti (1992: 7), että *translator is necessarily a betrayer*.
Tällä hän tarkoittaa sitä, että vaikka sanoille olisi esimerkiksi olemassa näennäiset käännökset eri
kielissä, ne eivät välttämättä kohtaa yksi yhteen. Tämä voi olla triviaalia, mutta paikoin myös on-
gelmallista kielenoppijalle, jos kohdekielestä ei tosiaankaan löydy sanaa tai rakennetta vastaamaan
sitä, jonka hän haluaa kohdekielellä ilmaista.

Käännettävyyden vaikeudesta kirjoittaa myös esimerkiksi Piotr Blumczynski artikkelis-
saan *Turning the tide: A critique of Natural Semantic Metalanguage from a translation studies*
perspective vuodelta 2013. Noin ylipäätään täydellistä käännettävyyttä pidetään käännöstieteessä
utopistisena, jolloin NSM:n kaltaisen yksityiskohtaista analyysiä painottavan teorian nousu on omi-
aan tarjoamaan selkeämpi kuva kääntämisen monimutkaisuudesta ja vastaamaan sen tuomiin haas-
teisiin.

Käännöstieteen lisäksi oleellisimpia NSM:n käyttökohteita ovat kuitenkin niin kutsu-
tut eksplikaatiot tai selitykset, joissa reduktiivista parafrasaa käytännössä hyödynnetään jonkin sa-
nan tai käsitteen selittämässä. Esitän eksplikaatiosta käytännön esimerkin. Seuraavassa on suora
lainaus siitä, miten Tissari ja Vanhatalo (2017) esittävät verbin *to promise* – *luvata* suomenkielisen
käännöksen, joka, kuten em. kuvailevat, koostuu kolmesta osasta: *kehyksestä* (lexico-syntactic
frame), *lausumasta* (dictum) ja *näkyvästä mentaalisesta tilasta* (apparent mental state). *Kehys* kuvaa

verbin olevan vaikuttamaan pyrkivä puheaktiverbi, *lausuma* sisältää sen, mitä agentti sanoo ja *näkyvä mentaalinen tila* kertoo sen, mitä agentti ajattelee sanoessaan näin. Esitän aluksi em. lainaaman alkuperäisen englanninkielisen eksplikaation (Goddard 2011: 142):

lexico-syntactic frame

a) someone X said something to someone else Y at that time this someone wanted something to happen because of it

dictum

b) this someone said something like this:

'I want you to know that I will do this (A) after this time I want you to know that I think about it like this now: 'I can't not do it' I know that if I don't do it after I say this, people can think something bad about me because of it'

apparent mental state

c) this someone said it like someone can say something like this to someone else when this someone thinks like this:

'I know that this someone
wants me to do it
I know that this someone can
think that I will not do it
I don't want this someone to
think like this'

leksikaalis-syntaktinen kehys

a) joku X sanoi silloin jotain jollekulle Y koska hän X tahtoi jotain tapahtuvan

lausuma

b) hän sanoi jotain tällaista:

'minä tahdon sinun tietävän, että minä teen tämän (A) tämän ajan jälkeen. minä tahdon sinun tietävän, että minä ajattelen tästä nyt näin: 'minä en voi olla tekemättä tätä' minä tiedän, että jos minä en tee tätä sen jälkeen, kun sanon tämän, ihmiset voivat ajatella minusta sen takia jotain pahaa'

näkyvä mentaalinen tila

c) hän sanoi tämän kuten joku voi sanoa toiselle ihmiselle kun hän ajattelee näin:

’minä tiedän, että tämä ihminen tahtoo minun tekevän tämän
minä tiedän, että tämä ihminen voi ajatella, että minä en tee sitä
minä en tahdo tämän ihmisen ajattelevan näin’

Kuten esimerkistä käy ilmi, eksplikaatio muistuttaa hyvin läheisesti normaalia kieltä: Se noudattaa kielen kieliopillisia sääntöjä ja lainalaisuuksia, se sisältää sellaisia kielen piirteitä, kuten vaikkapa lauseenvastikkeita, mutta samalla se vaikuttaa etäiseltä, ja ehkä hankalasti ymmärrettävältä. Tämä mahdollinen hankala ymmärrettävyys onkin yksi niistä NSM:n puolista, jotka ovat kirvoittaneet kritiikkiä. Pureudun kritiikkiin tarkemmin seuraavassa luvussa.

2.3 Kritiikkiä NSM:ää kohtaan

Natural Semantic Metalanguage -teorian piirissä on monia osa-alueita, jotka ovat saaneet osakseen runsastakin kritiikkiä. Kritisoitaviksi ovat joutuneet niin NSM:n perusajatuksot NSM:n pohjalla, ydinsanojen muodostamisperusteet, reduktiivistien parafrasien käytön järkevyys, primitiivien olemassaolo kuin yksittäiset alkusanat ja eksplikaatiotkin. On kuitenkin syytä huomata, että NSM on niittänyt myös ylistystä, ja sen sisällä tehtävää tutkimusta pidetään arvokkaana niin merkityksentutkimukselle kuin kielitypologiallekin. Koska NSM on pro gradu -tutkielmani ehdottomassa keskiössä, on mielestäni ensiarvoisen tärkeää pystyä tarkastelemaan sen paikkaa koko semantiikan tutkimuksen kentässä. Tämän katson tapahtuvan tehokkaasti kritiikin kuvailemisen kautta.

Tässä luvussa esittelen erityisesti Dirk Geeraertsin, Nick Riemerin, Piotr Blumczynskin, Maria Koptjevskaja-Tammin sekä Inger Ahgrenin ajatuksia, rinnallaan luonnollisesti NSM:n kahden suuren Anna Wierzbickan sekä Cliff Goddardin ajatukset. En väitä katsaukseni olevan kattava, mutta toivon, että onnistun siinä jotenkuten käsittelemään yleisimmin NSM:ssä kritisoituja aiheita.

Mahdollisesti eniten kritiikkiä kerääviä aiheita NSM:ssä on siihen useasti sisältyvä absoluuttisuuden ajatus. NSM esimerkiksi painottaa alkusanojen sisältyvän kaikkiin maailman kieliin. Sama absoluuttisuuden näkökulma esiintyy myös aukottomien parafrasien muodostamisessa. Itse katson, että NSM:ää, kuten kaikkea muutakin, tulee lähestyä hermeneuttisen periaatteen kautta: Aikaisempi

ymmärrys korjautuu ja täydentyy ymmärryksen kasvaessa; hypoteesi tarkentuu havaintojen perusteella. Tästä syystä ajatus ehdottomuudesta tulisi tosiaan hylätä, tai sitä tulisi ainakin kohdella mallina siitä, mihin tutkimus tähtää, ei siitä, mitä se on. Tätä voi verrata numeerisessa analyysissä käytettävään Newtonin menetelmään: Ensin suoritetaan alkuarvaus, jolla pyritään etsimään haluttu nollakohta. Sitten otetaan tämä alkuarvaus seuraavan laskutoimitukset lähtökohdaksi, jolloin saadaan alkuarvausta tarkempi arvio. Tämä tarkempi arvio otetaan seuraavaksi lähtökohdaksi, ja toimenpide toistetaan; eli suoritetaan iterointi. Näin lähestytään alati haluttua nollakohtaa.

Lisäksi haluan erikseen mainita periaatteen, jota kutsutaan nimellä *Strong Lexicalization Hypothesis*, ja johon viitataan tästedes akronyymillä SLH. SLH:n mukaan kaikki NSM:n alkusanat esiintyvät kaikissa maailman kielissä leksikaalistuineina, eli selkeinä sanoina tai rakenteina. Kuten Goddard itse esittää teoksessaan *Semantic and Lexical Universals: Theory and empirical findings* (1994), NSM:n seitsemäs periaate kuuluu: ”Strong Lexicalization Hypothesis. Every semantically primitive meaning can be expressed through a distinct word, morpheme or a fixed phrase in every language” Tähän absoluuttisuuteen pureutuu erityisesti Riemer artikkelissaan *Reductive Paraphrase and Meaning: A Critique of Wierzbickian Semantics* (2006). Jürgen Bohnemeyer (2003) puolestaan kuvaa SLH:ta sanoin: ”This is probably the most interesting hypothesis generated by the NSM program, just simply because it is the least likely one to be borne out.”

Koostan luvun kolmesta kritiikkiin pureutuvasta alaluvusta. Etenen seuraavasti: Liikun tyvestä kohti latvaa, tai ytimeistä kohti kuorta. Ensiksi esittelen, millaista keskustelua on käyty NSM:n perusrakenteisiin liittyvistä kysymyksistä: Teen tämän pureutumalla viiteen kysymykseen, jotka keskittyvät sellaisiin aihepiireihin kuin sanan yhtenäinen merkitys, pragmatiikan ja semantiikan erot, kieliyhteisö ja sen merkitys, täydellinen kääntäminen sekä ajattelun ja kielen suhde. Tässä luvussa en siis niinkään keskity itse NSM:ään kuin ajatuksiin, joiden päälle se on rakennettu.

Toisen alaluvun keskiössä on reduktiivinen parafraasi metodina, sen yhteys sanan merkitykseen sekä sen mekaaninen toteuttaminen. Siirryn siis luvussa käsittelemää NSM:n kulmakivien sijaan sen leipätyötä; sitä tutkimusta, mitä NSM:n avulla on tehty sitä käyttäen.

Kolmannessa ja viimeisessä alaluvussa esitän huomioita, joita on tehty erityisesti SLH:ta vastaan. Olen päättänyt erottaa tämän yhden NSM:n perusperiaatteen kritisoinnin omaksi luvukseksi, sillä se on kerännyt verrattain huomattavan määrän kritiikkiä. Tämän lisäksi käsittelen luvussa huomioita, joita on tehty liittyen semanttisten primitiivien olemassaoloon ja olemukseen.

Tässä luvussa en juuri pureudu yksittäisiin eksplikaatioihin tai alkusanoihin muutamia esimerkkejä lukuun ottamatta. Katson, että niiden kritisointi on hyödyllistä ja usein rakentavaa,

mutta usein teorian sisäistä toimintaa. Goddard (1998a, 1998b) itseasiassa peräänkuuluttaakin eksplikaatioiden ja alkusanojen tarkempaa analysointia, sillä se rakentaa teoriaa eteenpäin. Toisaalta, Goddard (1998b) toteaa kritiikin usein sijoittuvan kategoriaan, jota hän kutsuu nimellä *appeal to bad precedent*. Tällä hän tarkoittaa sitä, että huomion kiinnittäminen lähtökohtaisesti ongelmallisiin tapauksiin ei ole yleistettävissä koko teoriaa ansiokkaasti kuvaavaksi kritiikiksi.

Kiinnitän kyllä huomiota yksittäisten suomenkielisten alkusanojen olemuksiin työnsäni runsaastikin, mutta sen lisäksi, mitä olen jo NSM:n suomenkielistä käännöstä käsittelevässä luvussa 2.1 tehnyt, käsittelen aihetta lisää tutkielmani menetelmäluvussa 5, sillä pohjaan huomioni työni aineistoon.

Kirjoittaessani pyrin tuomaan esiin myös NSM:n teoreetikkojen näkemyksiä esittämäni kritiikkiin sikäli kuin mahdollista. Kuten Ulla Vanhatalo käymässämme sähköpostikeskustelussa mainitsee, osuvin kritiikki saataisiin aikaiseksi esittämällä muiden semanttisten teorioiden huomioita rinnan NSM:n kanssa. Mainittakoon tosin, että Riemer (2006) ottaa tällaiseen vertailevaan kritisointiin sellaisen kannan, että pyrkimällä olemaan *ajattelun kieli*, NSM:n asettaa itsensä asteittaisen vertailun ulkopuolelle. Toisin sanoen: Jos NSM:n tavoite on olla *tyhjentävä* (*exhaustive*), sen saaman kritiikin on oltava yhtä lailla tyhjentävää. Ei ole siis hänen mukaansa mielekästä vertailla, mikä semanttinen teoria on tehokkain missäkin tilanteessa, jos yksi näistä ilmoittaa olevansa absoluuttisesti paras.

Mitä tulee kritiikin luonteeseen, Goddard (1998b) toivoo itsekkin sen perustuvan loogisille ja perustelluille väitteille esimerkiksi siitä, miten joku primitiivi ei tosiaankaan kuulu jonkin tietyn kielen leksikkoon, tai miten jokin primitiivi voidaan hajottaa vielä pienemmiksi osiksi määrittelemällä se. Näillä molemmilla keinoilla siis todistetaan jompikumpi NSM:n primitiivien perusolettamuksista vääräksi kyseisen sanan kohdalla: Sanan on sisällyttävä kaikkien kielten leksikkoon, eikä sitä tule voida hajottaa enää pienemmiksi osiksi.

Esitän esimerkin, joka tiivistää joitakuuta NSM:nkin kohtaamia ongelmia: Betty Birner (<https://www.linguisticsociety.org/content/does-language-i-speak-influence-way-i-think> 6.2.2018) havainnollistaa näkemystään sanoista ja niiden merkityksistä esimerkillä lapsesta, joka oppii sanan *koira*: Ensimmäinen lapsi oppii, että perheen bernhardilainen on *koira*. Tämän jälkeen hän rakentaa mielessään koiran kategorian, johon ajan myötä lukeutuu kuulumaan myös naapurin chihuahua, mutta josta karsiutuu pois lehmä. Tämä kategoria jatkaa laajenemistaan, ja se sen rajat tarkentuvat ja tulevat selvemmiksi; lapsi oppii yhä paremmin tuntemaan, mikä on koira ja mikä ei.

Tämä ajatus on oleellinen kahdesta syystä: Ensinnäkin se havainnollistaa sitä, että kategoria, jota kutsutaan sanaksi, voi saada alati uusia merkityksiä. Näin ollen se ei ole milloinkaan

valmis, joten tuntuu mahdottomalta kehittää sille täysin paikkansa pitävää määritelmää. Toisekseen ei ole lainkaan selvää, milloin lapsen mielessä muodostuva kategoria saavuttaa sellaisen muodon, että sitä voi käyttää tehokkaasti kielessä. Lapsi voi käyttää sanaa *koira* väärin kuvaamaan naapurin kissaa, koska hän ei vielä tiedä sanaa tuolle eläimelle. *Koiran* kategoria ei siis saavuta lapsen mielessä täyttä mittaansa, mikä herättää kysymyksen: Kenen mielessä kategoria on niin täysi, että sitä voidaan alkaa määritellä? Tämä liittyy sekä semantiikan ja pragmatiikan eroihin että kieliyhteisön kokemukseen sanan merkityksestä. Näistä molemmista lisää myöhemmin.

On mielestäni asiallista mainita jo tässä vaiheessa, että perustelen luvussa 3, joka käsittelee toisen kielen oppimista ja omaksumista sekä SLA-tutkimusta myös sen, miksi esittämäni kritiikki ei mielestäni vesitä NSM:n käyttämistä opetustarkoituksessa. Lähtökohtaisesti en siis katso, että vaikka NSM:ssä on kysymyksiä, joihin minäkään en ole löytänyt vastausta, sitä tulisi missään tapauksessa hylätä teoriana tai lähestymistapana. Päinvastoin, kritiikin myötä herää niin parannusehdotuksia, kysymyksiä, joihin kaipaankin selkeämpiä vastauksia kuin näkökulmiakin, joihin uutta tutkimusta voi peilata.

Vaikka Riemer (2006) peräänkuuluttaakin sitä, että koska NSM:n esittää olevansa aukoton, kritiikin tulisi hyvin pitkälti kumota sen esittämät väitteet, en itse katso asiaa näin absoluuttisena. NSM on teoria, siinä missä muutkin, ja kuten Dirk Geeraerts (2009) esittää, eräs kaikkein edistyneimmistä yrityksistä saavuttaa kielenulkoisen metakieli: ”(NSM is) the most advanced attempt in contemporary semantics to establish an inventory of universal primitive concepts.”

2.3.1 NSM:n kulmakivet

Tässä luvussa pyrin esittämään näkökulmia siihen, millä keinoin sanalle tai käsitteelle voidaan löytää selkeä ja tarkkarajainen merkitys, ja onko sen etsiminen ylipäätään mielekäästä. Luvun sisältö rakentuu pohtimalla viittä kysymystä:

- 1: Onko sanalla tai käsitteellä yhtenäistä ydintä tai merkitystä – entä alkusanalla?
- 2: Mitkä ovat pragmatiikan ja semantiikan erot merkitystä etsittäessä.
- 3: Jakaako koko kieliyhteisö saman ymmärryksen tietyn sanan tai käsitteen merkityksestä?
- 4: Onko täydellinen kääntäminen mahdollista?
- 5: Heijastavatko sanat ajatusta?

Ensimmäinen kysymys: Onko sanalla tai käsitteellä yhtenäistä ydintä tai merkitystä – entä alkusanalla?

Tämän kysymyksen osalta pyrin esittämään huomiota siihen, onko millä tahansa sanalla olemassa jonkinlaista ydintä tai ydinmerkitystä. Tämä pohdinta jättää kuitenkin kokonaan auki sen kysymyksen, onko alkusanalla tällaista merkitystä. Tähän kysymykseen olen pyrkinyt vastaamaan tutkielmani luvussa 2.3.3, sillä tällöin kysymys on paremmassa asiayhteydessä.

Tämä on elintärkeä kysymys, kun pohditaan NSM:n perusteita, sillä Wierzbicka itse nimeää *vähentämättömän ytimen* (*irreducible core*) toiseksi NSM:n peruspiilareista: “The NSM approach to linguistic description is based on two fundamental assumptions: first, that every language has an irreducible core that enables its speakers to understand all complex thoughts and utterances; second, that the irreducible cores of all languages match, so that we can speak, in effect, of the irreducible core of all languages, which in turns reflects the irreducible core of human thought”. (Wierzbicka 2010, 17).

Toisin sanoen Wierzbicka siis muotoilee yleisen mallin, joka toimii perustana NSM-alkusanojen olemassaololle ja niiden olemukselle vähentämättöminä primitiiveinä. Tämän ajatuksen takia on tärkeää pyrkiä selvittämään, ovatko nämä alkusanat todellisia, ja ovatko ne juuri niin purkamattomia kuin Wierzbicka esittää.

Ensimmäinen sudenkuoppa tämän kysymyksen käsittelyssä on nähdäkseni se, ajatellaanko sitä, onko millä tahansa sanalla yhtenäinen sisäinen merkitys vai puhutaanko siitä, koostuuko kieli tosiaan tällaisista merkityksen osista, joita kasaamalla voidaan muotoilla mikä tahansa käsite. Ajatus siitä, että mikä tahansa sana ylipäättään voidaan määritellä jakamalla se pienempiin osiin (ks. esim. Riemer, 2006) olettaa, että sanoilla tosiaan on tällainen yhtenäinen stabiili merkitys. Toisaalta, voidaan leikitellä ajatuksella, että sanalla on yhtenäinen merkitys, mutta sen tuominen esiin käyttämällä pienempiä osasia voi osoittautua ylitsepääsemättömän laajaksi tehtäväksi. Toisin sanoen, näistä kahdesta oletuksesta se, että merkitys voidaan kaivaa esiin yksinkertaisemmilla sanoilla olettaa todeksi sen, että tällainen merkitys on olemassa, mutta tämä ei välttämättä toimi toisin päin.

Piotr Blumczynski lähestyy aihetta käännöstieteen näkökulmasta. Hän rinnastaa artikkelissaan *Turning the Tide: A Critique of Natural Semantic Metalanguage from a Translation Studies Perspective* (2013) ajatuksen *vähentämättömästä ytimestä* Nidan ja Taberin (1969, 1982) luomaan käsitte-

seen *kernel* ("ydin"), jolla pyrittiin saavuttamaan hyvin samankaltainen ratkaisu: "Jokaisessa kielessä on olemassa puoli tusinaa tai tusina perusrakenteita, joista yksityiskohtaisemmat muodot rakentuvat." (Nida & Taber 1969, 1982)

Nida ja Taber (ibid.) esittävät myös, että *kernelien* olemassaoloa tärkeämpää on se tosiasia, että kaikki kielet kohtaavat juuri *kernelien* tasolla, ei niinkään hienostuneempien rakenteiden. Tämä tarkoittaa heidän mukaansa sitä, että ylimääräiset kieliopilliset piirteet voidaan karsia *kernelien* ympäriltä pois, jolloin jäljelle jää helposti käännettävä ydin. Ajatus on siis hyvin samankaltainen kuin NSM:ssä. On kuitenkin huomioitava, että ajatus *kerneleistä* ei ole kestänyt sen aikaansaamaa vastustusta.

Blumczynski siteeraa artikkelisaan (ibid.) Nidaa (1985): "We are no longer limited to the idea, that meaning is centered in words or even in grammatical distinctions. Everything in language from sound symbolism to complex rhetorical structures carries meaning." Tästä Blumczynski vetääkin johtopäätöksen, että kokonaisen sanan tai käsitteen merkitys ei ole *ytimen* (kernel) piirre, vaan myös *kuoren*. Näin ollen ei siis ole mitenkään itsestään selvää, että merkitys voitaisiin hajottaa pienempiin komponentteihin, primitiiveihin, täysin tyhjentyvästi ottamatta huomioon sanan käyttöä eri tarkoituksissa. Tästä päästään seuraavaan kysymykseen.

Toinen kysymys: Mitkä ovat pragmatiikan ja semantiikan erot merkitystä etsittäessä?

Cliff Goddard ottaa kirjassaan (1998a: 6) selkeän kannan siihen, mitä merkitys on, tai vielä tarkemmin, mitä se ei ole. Ensinnäkin, hän painottaa, että merkitys ei ole tieteellinen totuus. Hän lainaa amerikkalaista lingvistiä Leonard Bloomfieldiä, joka painottaa tieteelliseltä tiedolta vaikuttavaa tarkkuutta vain joidenkin sanojen määrittelyssä, mutta sulkee tällaisen määrittelyn ulkopuolelle esimerkiksi sellaiset abstraktit käsitteet kuin *rakkaus* ja *viha*.

Toinen positio, jonka Goddard ottaa, on erityisesti suhteessa pragmatiikkaan: "*Meaning is not 'use'*" (Ibid.: 6). Hän perustaa ajatuksen merkityksestä sanan käyttönä Wittgensteiniin, ja myöntää itsekin, että sanan merkitys tosiaankin on kytköksissä sanan käyttöön. Hän kuitenkin painottaa, että tällaisen yksioikoisen ajatusmallin käyttö on hylättävä, tai sitä on ainakin muokattava rajusti.

Näillä lähtökohdilla Goddard siis asettaa ne puitteet, miten NSM:n piirissä tulee suhtautua merkitykseen ja pragmatiikkaan. Olisi erikoista, jos NSM:ssä hyväksyttäisiin ajatus merkityksestä täysin käytöstä riippuvana vaihtelevana entiteettinä, sillä saattaisi jopa vetää pohjan pois koko NSM:n teoreettiselta perustalta. Goddardin positio on siis oletetun kaltainen.

Esa Itkonen esittää teoksessaan *Maailman kielten erilaisuus ja samuus Osa 1* (2008) hieman vastavanlaisen näkemyksen, erityisesti siinä, mitä tulee Wittgensteinilaiseen psykologistiseen merkitysteoriaan, kuten Itkonen sitä nimittää. Itkosella on kuitenkin varsin selkeä malli siitä, mikä tosiaan on semantiikan ja pragmatiikan rajapinta: Hän käyttää esimerkkinä lausetta *Tulen luoksesi keskiyöllä*. Lauseen faktat ovat selvät, mutta sen merkitys tarkentuu, jos sitä käyttää rakastunut Romeo Juliaalleen. Yhtä lailla merkitys tarkentuu, jos vampyyri ilmoittaa saman tulevalle uhrilleen.

Tähän esimerkkiin vedoten Itkonen esittääkin mallin, jossa semantiikan ja pragmatiikan rajapinta noudattelee de Saussuren (1962 [1916]) klassista mallia *languesta* ja *parolesta*. Näin ollen olisi tosiaan olemassa merkitys, joka olisi yhtenevä *languen* kanssa. Sen sijaan ilmauksen käyttö ei automaattisesti olisi yhtä kuin merkitys, vaan se olisi tämän merkityksen *parole*-akti. Hän kutsuukin semantiikan tason merkitystä ”jäätäneeksi teoksi”, ja pragmatiikan tehtävää sen ”sulattamiseksi” (ks. Itkonen 1983).

Erityisen tärkeää on silti huomioda, että Itkonen (ibid.) painottaa semantiikan ja pragmatiikan suhteen dynaamisuutta. Tilannekohtainen merkitys voi tosiaan konventionaalistua, ja tämä nousu *languesta paroleen* olisikin ominainen kielen muutoksen piirre.

Haluan nostaa esiin pari esimerkkiä, joilla havainnollistaa semantiikan ja pragmatiikan rajapinnan konstikkuutta suhteessa NSM:ään. Sanat, joiden voidaan ajatella olevan vastineita NSM-alkusanoille tietyissä kielissä saattavat olla latautuneita konnotaatioiden tai käyttöympäristöjen perusteella: Kuten Vanhatalo, Tissari ja Idström (2014) toteavat, englanninkielisen alkusanan WANT suomenkieliseksi käännökseksi on pohdittu sanoja HALUTA ja TAHTOA. Heidän mukaansa kuitenkin HALUTA sisältää jotain alkukantaista tai jopa seksuaalista. Voitaisiin väittää, että myöskään TAHTOA, jonka käyttöön em. alun perin päätyivät, ei ole täysin vailla kasautuneita merkityksiä. Voidaan esimerkiksi huomioda, että sana *tahdon* esiintyy suomenkielisessä hääseremoniassa toisin kuin englanninkielinen sana *want*. Jos sanaa käytetään puhunnassa samanmallisena – *tahdon* – voi se jopa linkittyä hääseremoniaan. On syytä huomioda, että Vanhatalo itse kertoo (sähköpostikeskustelu) olevansa nykyään taipuvaisempi pitämään sanaa HALUA englanninkielisen alkusanat WANT suomenkielisenä vastineena.

Jos sanoilla on jonkinlainen ydin, jonka myös koko kieliyhteisö tiedostaa ja hyväksyy omakseen, tämä käyttötarkoitus sisältyy suomenkieliseen sanaan *tahtoa* kiteytyneenä, jolloin se vaikuttaa taustalla myös sanaa käytettäessä. Kyseessä ei välttämättä ole aivan Itkonen kuvaama siirtyminen *languesta paroleen*, mutta ei myöskään työkalun ja käyttökohteen kaltainen ei-riippuvainen sidos: Jakoavainta voi käyttää varasana, mutta se ei tee siitä vasaraa. Tämän rajapinnan epäselvyys kuvastaa mielestäni semantiikan ja pragmatiikan eron monimutkaisuutta.

Jos puolestaan esitetään, että tämä aiemmin kuvattu käyttötarkoitus ei sisälly sanaan *tahtoa* ainakaan kaikissa sen käyttöyhteyksissä, on katsoakseni perusteltua väittää, että sanoihin kasautuu ja kiteytyy erilaisia merkityksiä niiden käyttötilanteista riippuen, mikä asettaa ajatuksen täydellisesti määriteltävästä ytimeistä ajoittain kyseenalaiseksi.

Tämä huomio johtaa kysymykseen siitä, milloin pragmatiikasta – sanan käyttötarkoituksista ja -ympäristöistä – tulee osa sanan olemusta? Esimerkkinä toimii edelleen Vanhatalon, Tissarin ja Idströmin (2014) esimerkki sanasta *haluta*. Se saa aikaan erilaisen merkityksen riippuen sen käyttöympäristöstä. Sanalla *haluta* on ainakin kaksi rektiota: Jos sen objekti on akkusatiivissa tai genetiivimuotoisessa akkusatiivissa, sana ei välttämättä kanna minkäänlaista seksuaalista sävyä: *Minä haluan sinut joukkueeseeni*. Jos objekti sen sijaan on partitiivissa ja sattuu olemaan kohteena sopiva, seksuaalinen sävy on ilmeinen: *Minä haluan sinua*. Sen sijaan, vaikka objekti olisikin partitiivissa, mutta kohde itsessään ei ole sopiva, seksuaalinen sävy ei tule oikeastaan kysymykseenkään: *Minä haluan olutta*.

Kolmas kysymys: Jakaako koko kieliyhteisö saman ymmärryksen tietyn sanan tai käsitteen merkityksestä?

Kysymys on relevantti siksi, että jos kaikki, jotka tiettyä sanaa käyttävät, tiedostavat joko eksplisiitista tai implisiittisesti sen ydinmerkityksen, tämä ydinmerkitys voidaan kiistatta abstrahoida esiin, eli siitä voidaan muodostaa eksplikaatio käyttämällä reduktiivista parafrasaa. Lisäksi tällöin ei ole riskiä siitä, että kukaan kielen käyttäjistä käyttäisi sanaa yhteyksissä, joihin se ei kuulu. Toisaalta, jos voidaan olettaa, että kieliyhteisö ymmärtää tietyn alkusanan samalla tavalla, voidaan todeta, että kyseistä alkusanaa käyttämällä muodostettu eksplikaatio ei sisällä vaihtelua merkityksessään. Kääntäen, jos kieliyhteisö ei jaa ymmärrystä alkusanan merkityksestä, on turha olettaa sitä käyttäen muodostetun eksplikaation olevan paikkansapitävä.

Esitän esimerkkinä sanan *asia*, joka kuuluu myös Vanhatalon, Tissarin ja Idströmin (2014) listaan suomenkielisistä NSM-alkusanoista, ja jonka vastine englannin kielessä on *THING*. Sanalla *asia* tarkoitetaan Kielitoimiston sanakirjan (<https://www.kielitoimistonsanakirja.fi/net-mot.exe?motportal=80> 6.2.2018) mukaan ”toiminnasta, tapahtumista, olotilasta, ominaisuuksista tms. hahmottuva abstraktia kokonaisuutta”. Arkikielessä on kuitenkin mahdollista käyttää sanaa *asia* laajemmassa merkityksessä kuvaamaan myös sellaisia kohteita, joita kuvailtaessa olisi mielekkäämpää käyttää sanaa *esine*, joka puolestaan tarkoittaa Kielitoimiston sanakirjan (ibid. 2.6.2018)

mukaan ”määrätarkoitukseen käytettävää valmistetta, tavaraa tai välinettä”. Tästä huomiosta minulla ei valitettavasti ole kirjallista lähdettä, sillä kyseessä on itse tekemäni huomio. Mainitsen tosin, että myös Tampereen yliopiston suomen kielen professori Anneli Pajunen teki saman huomion luennollaan Tampereen yliopistossa syksyllä 2017.

Näin ollen sanan *asia* merkitys on ainakin joissain konteksteissa laajentunut, eikä ole enää yksiselitteistä, missä kontekstissa sitä tulee mitenkään käyttää. Sanoja voi tietenkin käyttää myös ns. tahallisesti väärin, jos halutaan saavuttaa puhunnalla jokin uudenlainen vaikutelma, mutta se, milloin tällaiset tarkoitukset ovat selkeitä ja milloin eivät, ei aina ole itsestään selvää. Yksittäinen käyttäjä ei siis omista sanan merkitystä; kuka tahansa voi päättää käyttää sanaa yhteydessä, jossa sitä ei tavanomaisesti käytetä joko vahingossa tai jossain muussa tarkoituksessa, esimerkiksi huumorimielellä.

Katson myös, että sanojen merkityksen pohtiminen kuvastaa omalla tavallaan sitä, etteivät niiden merkitykset ja käyttötarkoitukset ole aivan niin selkeitä, että tohtisimme väittää kieliyhteisön jakavan ymmärryksen jonkin sanan merkityksestä. Mielestäni tämä koskee niin alkusanoja kuin monimutkaisempiakin käsitteitä: Jos alkusanaksi ehdotettu sana osoittautuuakin niin monitulkintaiseksi, että sen merkitystä ei voida yksiselitteisesti hyväksyä, on pohdittava sitä, onko kyseinen sana alkusana ylipäänsä. Jos siis joudutaan erikseen miettimään sitä, mitä sana *asia* tarkoittaa, ei voitane olla täysin yksimielisiä kieliyhteisön yhteisestä ymmärryksestä ja siitä, että tämä alkusana tosiaan on yksiselitteinen ja hajottamaton. Tällöin on mahdollista, että kyseistä sanaa käyttämällä muodostetun eksplikaation tarkkuus vaarantuu. Paneudun aiheeseen alkusanakohtaisesti tarkemmin tutkielmasani.

Jos kaikki sanan sisältämät merkitykset siis eivät sisälly jokaisen yhteisön jäsenen käsitykseen sanasta, millä perusteella päätetään, mikä on oikea tulkinta sanan merkityksestä? Merkitys on mahdollisesti jonkin yhteisön yleisesti tunnustama, ja tämän yhteisön mielipiteen perusteella sanalle on annettu normatiivinen merkitys, joka on kirjoitettu esimerkiksi sanakirjaan. Näin ollen sanan merkitys voi olla vakiintunut sen käytön myötä tietynlaiseksi, jolloin se on rakentunut pienemmistä osista tai muuten muotoutunut. Tällöin voidaan myös ajatella, että mikä on rakennettu, voidaan myös purkaa.

Nämä asia huomioon ottaen on mielestäni relevanttia kysyä, miten päätetään, kenellä on oikeus luoda sanalle kattava eksplikaatio NSM:ää käyttäen. Jos tämän eksplikaation kehittäminen olisi kieliyhteisössä avoin tehtävä, johon kaikki voisivat osallistua, olisi teoriassa mahdollista, että sana saisi myös sellaisia merkityksiä, joita enemmistö kieliyhteisöstä ei tunnista. Olisiko tällöin siis nojattava

enemmistön valtaan, vai otettava huomioon kunkin määrittelijän kompetenssi kielenkäytön saralla. Tällöin olisi puolestaan riskinä se, että sanalta jää puuttumaan sellaisia ominaisuuksia, joita joku toinen sille antaisi.

Lisäksi haluan edelleen painottaa, että kieliyhteisön ymmärrys sanoista ulottuu myös alkusanoihin asti. Jos voitaisiin todeta, että lähtökohtaisesti yhteisön sisällä vallitsee yhteisymmärrys NSM:n alkusanoihin kuuluvien sanojen olemuksesta ja merkityksestä, voitaisiin melko turvallisesti lähteä etenemään reduktiivisen parafraasin kehittämisessä kyseisten sanojen perusteella, jakoi yhteisö ymmärryksen monimutkaisemmasta sanasta tai ei. Mutta jos taas asetetaan kyseenalaiseksi se, ymmärretäänkö tietty alkusana tismalleen samalla tavalla yhteisön sisällä, myös reduktiivisen parafraasin käyttö saattaa hankaloitua huomattavasti.

Painotan vielä, että jos käsitteelle halutaan muodostaa aukoton eksplikaatio, täytyy vallita yhteisymmärrys siitä, että eksplikaation kirjoittamiseen käytettyihin alkusanoihin ei sisälly monitulkintaisuutta. Jos tämä oletus ei täyty, eksplikaatio jättää huomattavasti vaihtelunvaraa itsensä ja kyseessä olevan määriteltävän käsitteen välille.

Neljäs kysymys: Onko täydellinen kääntäminen mahdollista?

Kysymys on oleellinen siitä syystä, että esimerkiksi Vanhatalo (luentosarja 2014) esittää NSM:n yhden käyttötarkoituksen olevan kulttuurispesifien käsitteiden kääntäminen vaivatta kielestä toiseen. Tämä tapahtuu siis muodostamalla monimutkaisesta, kulttuurispesifistä käsitteestä, kuten vaikkapa suomen kielen sanasta *sisu* parafraasi, joka sisältää koko sanan merkityksen alkusanoin auki kirjoitettuna. Näin kyseinen sana voidaan siis siirtää kielestä toiseen ilman pelkoa merkityksen muuttumisesta: Tarkoituksena on siirtää vain kyseisen sanan merkitykseen koodautunut tieto, ei enempää, ei vähempää. Täydellistä kääntämistä tavataan kutsua luonnolliseksi ekvivalenssiksi (Blumczynski, 2013.)

Koko Blumczynskin artikkeli (2013) tosiaankin pohjaa nimenomaan käänntieteelliseen näkökulmaan, vaikka olenkin hänen ajatuksiaan lainannut jo aiemmin. Tämä näkökulma painottuu erityisesti, kun otetaan huomioon Goddardin ja Wierzbickan ajatukset alkusanojen ja reduktiivisen parafrasien käytöstä täydellistä käännettävyyttä tavoiteltaessa.

Goddardin ja Wierzbickan (1995) mukaan sanan eksplikaatio voidaan *siirtää* kielestä toiseen ilman menetyksiä tai vaihtelua merkityksessä, jos se tehdään täysin universaalien semanttisten primitiivien eli tai niiden kielellisten ilmentyminen, alkusanojen avulla. Täydellisen ytimen löytäminen ja täydellinen käännettävyys kulkevat siis käsi kädessä, sillä toisaalta ytimen löytäminen ja

olemassaolo ovat NSM:n oleelliset taustatekijät ja toisaalta tämän ytimen välittäminen tai kääntäminen toiselle kielellä on NSM:n perustavanlaatuinen lähtökohta.

Blumczynski (2013) esittää kuitenkin näkökulmia siihen, miksi tämä on hankalaa. Hänen tärkein teesinsä on se, että täydellinen semanttinen ekvivalenssi toimii vain suljetuissa systeemeissä, joissa sanoilta on riisuttu täysin niiden polyseemiset ja homonyymiset elementit, puhumattakaan sanojen metaforisesta olemuksesta. Vanhatalo huomauttaa sähköpostikeskustelussa, että tähän NSM:n puolelta vastaavat esimerkiksi ns. kanoniset lausekkeet. Näillä tarkoitetaan listaa sellaista asiayhteyksistä ja esiintymistä, joissa alkusanojen katsotaan esiintyvän. Lähestyn kanonisia lausekkeitä tarkemmin tutkielmani menetelmäluvussa 5.

Aikanaan Nidan ja Taberin (1969, 1987) esittäessä kernel-mallinsa siitä kiinnostuttiin myös käännöstieteen saralla (Blumczynski 2013) – onhan kyseessä teoria, joka olisi ollut hyvinkin hedelmällinen käännöstiedettä ajatellen. Ei siis liene ihme, että myös NSM on herättänyt käännöstieteessä huomiota. Blumczynski (2013) kuitenkin siteeraa Peter Fawcettia (1997, 67), joka kuvaa *kernel*-mallin olevan ”monimutkainen, jopa epämääräinen, ja kiinnostunut saavuttamaan vain yhden näkökulman käännöstieteisiin” ja ”ratkaisevasti hyödytön käännöstieteen tarpeisiin”.

Wierzbicka (1992) myöntää itsekkin, että oleellinen kysymys käännöksiä käsiteltäessä ei välttämättä ole, josko jokin sana tai käsite on käännettävissä, vaan pikemminkin tulisi pohtia, missä määrin kääntäminen on mahdollista. Kääntäminen ei siis esiinny *päällä/pois*-tyylisenä polaarisenä toimenpiteenä, vaan lineaarisena prosessina, joka tähtää minimoimaan vääristymät, ylimääräiset merkitysisällöt ja puutteet: “The real question, then, is not whether meaning can be transferred from one language to another but to what extent it can be so transferred; not whether meaning is language-independent but to what extent it is. Or, to put it differently, to what extent languages are shaped by ‘human nature’ and to what extent they are shaped by culture.” (1992, 7) Tämä on tärkeää, sillä tällä lausunnolla Wierzbicka eliminoi muuttujan, johon kriitikon on kaikista helpoin tarttua: kääntämisen täydellinen onnistuminen.

Kuten edellisessäkin sitaatissa ilmenee, NSM:ään kuuluu oleellisesti ajatus kulttuurispesifisyydestä: Sana, joka esiintyy jossain kieliympäristössä, sisältää väistämättä merkityksiä, joita ei esiinny jossain toisessa kulttuuriympäristössä esiintyvässä sanassa, vaikka ne tarkoittaisivatkin näennäisesti samaa asiaa.

Wierzbicka (1997) havainnollistaa kulttuurispesifien sanojen luonnetta rinnastamalla englannin kielen sanan *freedom*, latinan sanan *libertas*, venäjän kielen sanan *svoboda* ja puolan sanan *wolnosc*. Wierzbickan mukaan nämä sanat ruumiillistavat eri käsitteitä, jotka heijastavat erilaisia kulttuurisia ajatuksia. Niiden täydellinen ymmärtäminen voidaan Wierzbickan mukaan taata

vain tulkitsemalla niitä omaa kielikulttuuriaan vasten, jolloin ne antavat myös viitteitä kyseisestä kielikulttuurista. NSM:n avulla voisi siis olla mahdollista purkaa näitä kulttuurisidonnaisia merkityksiä käännettäväksi toiselle kielelle.

Blumczynski tähdentää, että jotta voitaisiin todella kääntää sana toiselle kielelle, täytyisi myös varmistua siitä, miten ylimääräisiä tai puuttuvia merkityksiä voidaan mitata. Tämä pätee myös alkusanojen käännöksiin: Jotta voitaisiin olla varmoja siitä, että alkusanojen käännökset tosiaan ovat täysin ekvivalentteja toisiinsa nähden, tämä yhteneväisyys täytyisi jollain tavalla voida varmistaa.

Blumczynsky sanookin, että tällöin mukaan tarvittaisiin *tertium comparationista*, joka ajatuksena on peräisin Jeremy Mundaylta. Tarvittaisiin siis kolmatta osapuolta; mittakeppiä, jota vasten merkitystä voidaan verrata ilman, että se tehdään kummankaan verrattavana olevan kielen kautta. Tällaisen työvälineen käyttö nojaisi siis empiirisyyteen. Ongelmana on pääsy kielen metatasolle, *langueen*. Blumczynski katsoo, että *languen* tavoittelu voi lopultakin olla vain teoreettista: Puhuminen ja kielen käyttö vastaa aina *parolea*.

Myös Riemer (2006) sivuaa aihetta: ”Without such an independent determination the argument for the correctness of the NSM paraphrase is both stipulative and circular: we are asked to accept an NSM definition as a true representation of the meaning of a definiendum because it does not involve any addition to or loss from this meaning – because, in other words, it is a true representation of its meaning.

Kuten NSM:n nimestäkin ilmenee, NSM pyrkii olemaan tällainen metakieli, joka voisi toimia merkityserojen erittelijänä ilman kielisidonnaisuuksia. Wierzbicka (1991) esittääkin: ”To compare meanings expressed in different languages and different cultures, one needs a semantic metalanguage independent, in essence, of any particular language or culture – and yet accessible and open to interpretation through any language.” Tähän näkökulmaan perustuen NSM voisi toimia juurikin Blumczynskin peräänkuuluttamana *tertium comparationiksena*, mutta se törmää yhteen ongelmaan: Ennen kuin NSM voitaisiin ratifioida päteväksi mittakepiksi kääntämisen tarkkuutta arvioitaessa, sitä itseään täytyisi testata: Pitäisi kyetä tarkastelemaan, millaisia latautuneita merkityksiä NSM:n itsensä piirissä on havaittavissa. Kuten Goddard (1998a) esittää, tätä on hankala aukottomasti todistaa. Ainoa keino löytää esimerkiksi alkusanoista merkityseroja on pyrkiä määrittelemään alkusana itse reduktiivisen parafrasoin avulla.

NSM:n täytyisi näin ollen olla suure, joka nojaa johonkin muuttumattomaan, totuutena pidettyyn asiaan, kuten esimerkiksi metrin määritelmä SI-järjestelmän mukaan: Metri määritellään nykyään pituudeksi, jonka valo kulkee tyhjiössä 1/299 792 458 sekunnissa (Wikipedia 7.2.2018).

Tässä tullaan siis siihen huomioon, että NSM alkusanat ovat kielisidonnaisia: Niitä ei voi ilmaista ihmiskielen ulkoisella puhtaalla *languella*. Jotta nämä sanat voitaisiin ratifioida toimiviksi ja täydellisen riippumattomiksi verrokkisanoiksi, niiden riippumattomuus täytyisi voida todistaa. Toisin sanoen NSM täytyisi itse validoida *tertium comparationista* vasten, mutta koska tarkoituksena on testata NSM:ää juuri tuohon tarkoitukseen, asetelma muodostuu paradoksaaliseksi.

Vanhatalo (sähköpostikeskustelu) kuitenkin huomauttaa, että tämä heijastelee kielitieteen suurta paradoksia: Kohde ja menetelmä ovat sama, jolloin voidaan joko lopettaa koko kielen tutkimus kokonaan, tai pyrkiä minimoimaan ristiriidan aiheuttamat ongelmat, ja jatkaa tutkimusta.

Viides kysymys: Heijastavatko sanat ajatusta?

Kysymys kielen ja ajattelun suhteesta on kiinnostanut ihmisiä satoja vuosia. Tunnetuimpia esimerkkejä aiheesta lienee kuitenkin niin kutsuttu Sapir-Whorf-hypoteesi. Se on aksiooma, jonka perusteella ihmisen puhumalla kielellä ja hänen maailmankuvallaan on systemaattinen ja yksisuuntainen relaatiiosuhde, jossa henkilön käyttämä kieli ohjaa hänen ajatteluaan. Teoria on laajalti kritisoitu, mutta kokonaisuudessaan ihmisen ajattelun ja kielen kytköstä ei ole kyetty ratkaisemaan.

Kuten Betty Birner (<https://www.linguisticsociety.org/content/does-language-i-speak-influence-way-i-think>) asian ilmaisee, kyseessä on tavallaan muna-kana-kysymys: Enkö voi ajatella jotain, koska minulla ei ole kielitaitoa sen kuvaamiseen, vai puuttuko minulta kielitaito asian kuvaamiseen, koska en ajattele sitä koskaan?

Kuten jo johdannossa kuvasin, Birner (ibid.) havainnollistaa ajattelun ja kielen suhdetta lapsella, joka oppii sanan *koira*. Lapsi oppii sanan aikanaan, ja kun hän osaa sen, hän oppii lisää kohteista, jotka voi asettaa tähän kategoriaan. Hän oppii myös niistä kohteista, joita ei voi koiraksi laskea. Näin hän rakentaa itselleen käsityksen siitä, mikä on koira, mitä piirteitä siihen kuuluu ja minkälaiset entiteetit voivat kuulua koiran kategoriaan

Syy, miksi kysymystä ajattelun ja kielen suhteesta on oleellista käsitellä, on se, että Wierzbicka itse sanoo kielen heijastelevan ei niinkään maailmaa, vaan ihmisen käsitystä ja tulkintaa maailmasta: ”language doesn’t reflect reality directly: it reflects human conceptualisation, human interpretation of the world” (Wierzbicka 1992, 7).

Jos siis kieli on korkeintaan hyvin toimiva representaatio ympäröivästä maailmasta, miten voidaan otaksua, että tarkastelemalla kieltä voitaisiin luoda NSM:n etsimä *alphabet of thought*? Samaan tapaan kuin koira, ja lapsen käsitys koirasta Birnerin (internetlähde, 7.2.2018) esimerkissä, ajatus ja sen kielennoös eivät välttämättä milloinkaan kohtaa toisiaan. Ajatuksen pukeminen kielen asuun voi tarkentua, olla hyvin lähellä, mutta jotta se todella kuvaisi ajatusta, sen tulisi

olla täydellinen. Jotta se ei olisi täydellinen, siihen ei pitäisi voida lisätä enää mitään, mikä tuo siihen lisäinformaatiota. Kuitenkin, yhden ihmisen lienee mahdotonta hahmottaa jonkin entiteetin kokonaista täydellistä olemusta: Onko Suomessa asuvalle lapselle oleellista tähdentää, että koira voi vetää valjakkoa? Entä Italiassa? Mitä tapahtuu, kun koiraa jalostetaan, ja kehittyy uusia rotuja, jotka täytyy mahduttaa sanan *koira* kategoriaan?

Artikkelissaan *Turning the Tide: A Critique of Natural Semantic Metalanguage from a Translation Studies Perspective* (2013) Piotr Blumczynski kiinnittää huomionsa Wierzbickan teoksiinsa sisällyttämiin seuraaviin sitaatteihin:

- Benjamin Lee Whorf (1956): We dissect nature along lines laid down by our native languages'' (Wierzbicka 1997, 7)

- Edward Sapir (1949) "No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same world with different labels attached" (Wierzbicka 1992, 4)

- Wilhelm von Humboldt (1988): "There resides in each language a characteristic world-view." (Wierzbicka 2010, 4)

Näiden havaintojen perusteella voidaan sanoa, että NSM:n taustalla on jokseenkin selkeä kuva ajattelun ja kielen suhteesta; kysymyksestä, johon selkeän vastauksen esittäminen vaatisi tarkkaa tietoa ajattelun olemuksesta. Juuri tätä filosofista premissiä Piotr Blumczynski (2013) artikkelissaan pohdii.

Blumczynski tähdentää NSM:n painottavan sitä, että kieli ei heijasta ajatusta puhtaasti, vaan se on aina tulkintaa. Myös Blumczynski ottaa rinnalle De Saussuren käsitteet *langue* ja *parole*. Jos kieli on aina ajatusten tulkintaa, niin se luonnollisesti edustaa De Saussuren *parolea*, jolloin ajatus vastaa *languea*. Sanan tai käsitteen todellisen ytimen etsiminen tähtää siis *languen* löytämiseen, mikä on Blumczynskin mukaan ongelmallista: Kääntäminen tapahtuu lopulta aina *paroleiden* kesken, sillä kieli on edellisen perusteella aina *parolea*. Tämä altistaa merkityksen aina tilanteiselle vaihtelulle. Huomioitavaa on myös se, että kaikki selitykset, joita voimme tuottaa ovat sidottuina käyttämäämme kieleen; emme voi puhua metakieltä, vaan kaikki ilmaisut esitetään kielellä, joka on lähtökohtaisesti *parole*. Tämä on tärkeää myös NSM-alkusanoja pohdittaessa: Vaikka alku-

sana pyrkisi kuvaamaan kielen *langue* tasoa, se on silti olemassa vain jollakin tietyllä kielellä *parole*-esiintymänä. Toisin sanoen alkusana sisältyy aina jonkin kielen leksikkoon. NSM tähtää nimensä mukaisesti metakielen kehittämiseen, mutta siitä huolimatta jokainen NSM-alkusana esiintyy vain tietyssä kieliasussa. Alkusanat voidaan kääntää toisille kielille, mikä jo itsessään tuottaa aika ajoin ongelmia, mutta jos niillä halutaan kuvata ajattelun komponentteja, tulisi päästä täydellisesti irti sanojen mukanaan kantamista merkityksistä. Blumczynski (ibid.) painottaa: Meillä ei ole keinoa päästä tutkimaan käsitteen tai sanan olemusta kielen taa, vaan meidän täytyy tyytyä kuvailemaan sitä kielen sisäisin keinoin.

Dirk Geeraerts käsittelee teoksessaan (2009) laajasti eri semanttisia malleja, ja uhraa myös luvun malleille, joista hän käyttää yläkäsitettä *decompositiona approaches*. Näin mallien tarkoituksena on siis purkaa rakenteita tarkastellakseen, mistä ne koostuvat. Ajatus on sikäli epäholistinen, että siinä katsotaan kokonaisuuden olevan yhtä suuri kuin osiensa summa.

Itse purkavan metodin mielekkyyteen Geeraerts esittää kolme aforisminomaista lausetta, joilla hän pyrkii tiivistämään eri purkavaa metodia hyödyntävien mallien kannan: ”the mind is neat but the world is fuzzy”, ”conceptual knowledge is parsimonious, but perceptual knowledge is abundant” ja ”semantics is stable but pragmatics is flexible”. Hän esittää näistä eritoten ensimmäisen kuvaavan NSM:ää: Geeraertsin mukaan ajatus siitä, että mieli on järjestyksessä ja kontrolloitu, mutta maailma on lähtökohtaisesti sekava, kuuluu oleellisesti NSM:n peruskiveen.

Kuten D. W. Murray artikkelissaan *Human Emotions: Some Problems of Wierzbicka's Simples* (1988) Proustin sanoin esittää: ”We feel in one world, we think and name in another. Between the two we can set up a system of references, but we cannot fill the gap.” Tähän kuitenkin Vanhatalo (sähköpostikeskustelu) lisää, että aina voi yrittää; se on parempi vaihtoehto kuin vain nostaa kädet ilmaan.

2.3.2 Reduktiivisen parafraasin mielekkyys

Reduktiivinen parafraasi on metodi, jolla pyritään määrittelemään sana tai käsite nimensä mukaisesti jakamalla sen merkitys pienempiin osiin. Sen tarkoituksena on siis kirjoittaa sanan merkitys auki käyttäen yksinkertaisempia sanoja kuin määriteltävänä oleva sana itse. Yleensä NSM:n piirissä reduktiivista parafrasaa käytetään hyödyntämällä alkusanoja. Reduktiivista parafrasaa käyttämällä muovattua selitystä sanan merkityksestä kutsutaan eksplikaatioksi (*explication*) tai selitykseksi.

Lähtökohtaisesti reduktiivista parafraasia käytettäessä tulisi pitäytyä käyttämästä muita sanoja kuin NSM:n piiriin kuuluvia alkusanoja; semanttisia primitiivejä, joita ei tule voida enää hajottaa pienempiin osiin. Goddard (1998a) kuitenkin mainitsee, että tämä ei ole aina ehdottoman välttämätöntä: Ainoa ehdoton asia on se, että eksplikaatio täytyy koostaa semanttisesti yksinkertaisemmista osista kuin määriteltävänä oleva sana itse. Näin vältetään sirkulaation eli kehämääritelmän (*circulation*) riski.

Sirkulaatiolla eli kehämääritelmällä (ks. esim. Goddard 1998a: 61) tarkoitetaan sitä, että jokin sana voidaan määritellä käyttämällä sellaisia sanoja, jotka puolestaan voidaan määritellä käyttämällä alun perin määriteltyä sanaa. Jos käsitteen tai sanan määritelmässä esiintyy sirkulaatiota, sitä ei voi tyhjentävästi määritellä.

NSM-eksplikaatiot voivat myös sisältää niin kutsuttuja molekyylejä, joita merkitään alaindeksillä *m*, esimerkiksi *auto_m*. Tämä tarkoittaa, että näille sanoille on jo olemassa eksplikaatio, ja sitä ei tarvitse erikseen toistaa kyseisessä asiayhteydessä.

Dirk Geeraerts esittää teoksessaan *Theories of Lexical Semantics* (2009) perusteita sille, miksi komponentteihin nojaavien lähestymistapojen piirissä usein ajatellaan, että määrittelyn on päädyttävä lopulta elementteihin, joita itse ei voi enää määritellä käyttämällä yksinkertaisempia sanoja. Näitä elementtejä NSM:ssä vastaavat primitiivit, joita käyttämällä voidaan siis muodostaa sanan ydin. Suurimmaksi tekijäksi Geeraerts esittää kehämääritelmien välttämisen.

Esimerkiksi Goddard puolustaa vahvasti tällaisen reduktiivisen määrittelyn käyttöä. Artikkelissaan *Bad Arguments Against Semantic Primitives* (1998) hän lainaa Antoine Arnauldin ja Pierre Nicolen tunnettua tekstiä *Logic, or, the Art of Thinking* (1996 [1662]): ”... I say it would be impossible to define every word. For in order to define a word it is necessary to use other words designating the idea we want to connect to the idea being defined. And if we wished to define the words used to explain that word, we would need still others and so on to infinity. Consequently, we necessarily have to stop at primitive terms which are undefined.” (Arnauld & Nicole 1996[1662]: 64)

Nick Riemer (2006) lähestyy reduktiivista parafraasia siitä näkökulmasta, että edistääkö käsitteen purkaminen pienempiin osiin todella väistämättä sen ymmärrettävyyttä. Hän rinnastaa kaksi erilaista lähestymistapaa: Onko mieltä pyrkiä tekemään eksplikaatio, joka on luotu käyttäen NSM:n piirissä yleisesti hyväksytyjä primitiivejä lähestymättä asiaa tilannekohtaisesti, vai tulisiko ensisijaisesti pyrkiä käyttämään sellaisia sanoja, joiden voidaan olettaa olevan vastaanottajalle jo valmiiksi tuttuja.

Kysymys on sikäli pätevä, että se ottaa suoraan kantaa merkityksen välittämiseen ymmärtämisen mittarilla, mikä ei suinkaan ole automaattista, vaan riippuu aina yksilöiden positiosta. Tässä asiassa kuitenkin piilee mielestäni Riemerin kysymyksen sudenkuoppa: Jos asiaa todella lähestyttäisiin tapauskohtaisesti, millä perusteella yleinen malli edes luotaisiin? Tapauskohtaisenkin tarkastelunkin tulee perustua jollekin lähtökohdalle, ja mielestäni NSM tarjoaa tällaisen lähtökohdan monessa tapauksessa varsin tehokkaasti. NSM:n piirissä painotetaan myös, että sen primitiivit ilmaantuvat ihmisen kieleen hyvin varhaisessa vaiheessa (Goddard 2001), jolloin on mielestäni perusteltua katsoa niiden myös tulevan ymmärretyiksi.

Myös Piotr Blumczynski (2013) lähestyy reduktiivista parafraasia samasta näkökulmasta. Hän siteeraa Riemerää (2006): ”It only makes sense to believe in the existence of semantic primitives if we believe, that explanation proceeds via reduction to simpler elements”. Myös Blumczynski katsoo, että ymmärtääksemme käsitteen, ei ole välttämättä järkeä hajottaa sitä pienimpiin mahdollisiin osiin, vaan ainoastaan pienimpiin tuntemiimme. Osien ei siis tarvitsisi olla mahdollisimman yksinkertaisia, vaan mahdollisimman tunnettuja. On hyvä huomioida, että Blumczynskin lähtökohta on kuitenkin käänntietieteissä.

Blumczynskin huomiot, vaikkakin paikkinsapitävät, eivät mielestäni kumoa omaa vastaustani Riemerin kritiikkiin. Lisäksi Ulla Vanhatalo (sähköpostikeskustelu) huomauttaa, että juuri tähän asiaan NSM:n puitteissa vastaavat molekyylit: Ne ovat käsitteitä, joita voidaan käyttää eksplikaation luomisessa, mutta jotka ovat monimutkaisempia kuin alkusanat itsessään. Lisäksi niille on olemassa kattava eksplikaatio valmiiksi.

Myös Cliff Goddard kirjoittaa teoksessaan *Semantic Analysis: A Practical Introduction* (1998), että semanttisten primitiivien käyttäminen eksplikaatiossa ei ole välttämätöntä: Eksplikaatio voi sisältää myös primitiivejä monimutkaisempia termejä ja olla silti merkityksellinen.

Teoksessaan *Semantic Analysis: A practical Introduction* (1998) Goddard esittää reduktiivisen parafrasien ja sen avulla muodostetun eksplikaation lopulliseksi testausmenetelmäksi *korvattavuutta* (*substitutability*). Tämä tarkoittaa sitä, että eksplikaation tulee olla vaihdettavissa sanaan, josta se on tehty missä tahansa käyttöympäristössä vailla mitään muutosta merkityksessä. Luonnollisesti saman mekanismin tulee toimia myös toisin päin. Vielä tarkemmin asiasta on kirjoittanut Wierzbicka (1988): “Only if the definiens can be substituted for the definiendum without loss or addition of meaning (*salvo sensu*) in the original context (*in locum*) is it accepted as its correct analysis.”

Riemer (2006) esittää asiasta seuraavan kuvion:

1. Substitutability

Linguistic elements (x and y) can be substituted for element z

(‘unmarried’ + ‘male’ can be substituted for ‘bachelor’)

therefore

2. Identity

The meaning of z is identical to the composition of the elements (x and y).

(the meaning of bachelor is identical to the composition of the two elements

‘unmarried’ and ‘male’).

Hän kuitenkin vertaa käsitteitä *substitutability* ja *identity* seuraavasti: *Identity* kuvaa jonkin sisintä olemusta, kun taas *substitutability* kuvaa vain ekvivalenssia tietyssä tilanteessa, toisin sanoen sen roolia jossain määritellyssä kontekstissa. Jos siis, kuten Riemer esittää, oletetaan sanalla olevan kiinteä, semanttinen ydin, sen vertaaminen suoraan kyseisen sanan korvattavuuteen jossain rakenteessa vaikuttaa aivan eri asialta. Se, että tämä kyseinen elementti voidaan korvata yhdessä kontekstissa omalla eksplikaatiolla, kertoo siis vain sen, että näin voidaan toimia juuri tässä yhteydessä. Lähtökohtaisesti kyseessä on siis Riemerin mukaan kaksi eri käsitettä: määritelmä ja merkitys.

Blumczynski (2013) pohtii Wierzbickan lausuntoa siitä, että sanat ovat käsitteiden ruumiillistumia (Wierzbicka 1997: 153). Hän kysyykin, että voiko käsitteellä sitten olla kaksi tai useampia mahdollisia ruumiillistumia, vai onko olemassa vain yksi? Hän esittää kaksi mahdollista ongelmaa tähän ajatukseen viitaten: Jos käsitteellä voi olla vain yksi ruumiillistuma, eikö se saata itse sanan käytön huomattavan kyseenalaiseksi? Tällöinhän sanaa voisi käyttää vain tietyssä kontekstissa, ja sen käytön monipuolisuus laskisi huomattavasti. Jos käsitteellä puolestaan voi olla monta ruumiillistumaa, sen merkitys ei vaikuta olevan lainkaan niin tarkkarajainen, kuin NSM olettaisi sen olevan. Tällöin itse merkityksen saavuttaminen reduktiivisen parafrasien avulla hankaloituisi.

Mielestäni vastaus Blumczynskin ensimmäiseen ongelmaan tulee annetuksi edellisessä kappaleessa Riemerin (2006) ajatuksin: Merkitys ja käytettävyys voivat olla toisistaan täysin erillään olevia asioita, jolloin sanan käytön kyseenalaistaminen yhden ruumiillistuman perusteella sekoittaa kaksi käsitettä keskenään. Blumczynskin toinen ongelma puolestaan pohjaa syvälle NSM:n ja muidenkin semanttisten teorioiden peruspiireihin ja on varsin pätevä. Sen käsittely tässä uudestaan olisi kuitenkin hyödytöntä, sillä aihetta on käsitelty jo tutkielmani luvussa 2.3.1 kysymyksissä yksi ja kolme.

Blumczynski esittää kuitenkin mielestäni hyvin pätevän kannan liittyen siihen, kuinka hyvin NSM-eksplikaatiot todella vastaavat niitä käsitteitä, jotka niiden on tarkoitus kattaa. Hän antaa esimerkin Wierzbickan teoksesta (2010: 19-20), jossa verrataan englannin kielen sanaa *empirical* ja ranskan kielen sanaa *empirique*:

empirique (e.g. approach)

someone thinks like this (about something):

“I want to know some things about this

maybe I can know these things if I do some things

because of this, I want to do some things

I don’t know well what I can do

I don’t want to think about it for a long time

I know that before, when people did some things

they could know some things about things like this because of this

maybe the same will happen to me now”

it can be bad if someone thinks like this

empirical (e.g. approach)

someone thinks like this (about something):

“I want to know some things about this

I know that people can’t know things like this about something

if they don’t do things of some kinds to some things

if people do things of these kinds to some things, they can see some things because of this

at the same time some parts of their bodies can touch some things

after this, they can know some things because they have done these things

I want to do some things like this now”

it is good if someone thinks like this

Blumczynskin ensimmäinen huomio näihin eksplikaatioihin liittyy siihen, millaista kieltä ne käyttävät ja millaisia mielikuvia ne herättävät. Hän huomauttaa, että lause ” if people do things of

these kinds to some things, they can see some things because of this” ei herätä hänessä mitään skenaariota, johon sitä voisi poissulkevasti soveltaa. Tai kääntäen, se herättää niin paljon mahdollisuuksia, että on mahdoton valita yhtä ainoaa. Lause vaikuttaa siis Blumczynskin mukaan liian laajalta asetettavaksi kontekstiin. On mielestäni kuitenkin perusteltua huomauttaa, että eksplikaatiot toimivat kokonaisuutena. Yhden rivin irrottaminen eksplikaatiosta ei ole siis katsoakseni mielekäästä, ja on varsin ymmärrettävää, ettei se löydä vastinetta reaali maailmasta.

Huomio, jonka Blumczynskin ainakin implisiittisesti esittää on siis myös se, että nimensä mukaisesti *reduktiivinen* parafraasi toimii vain yhteen suuntaan: Se purkaa merkityksiä. Jos todella haluttaisiin saavuttaa sellainen taso, että mikä tahansa käsite olisi korvattavissa missä tahansa kontekstissa omalla eksplikaatiollaan, sen tulisi, ollakseen täysin ekvivalentti, toimia myös toisin päin: Eksplikaatiosta pitäisi siis voida päätellä, mitä käsitettä se kuvaa. En väitä, etteikö tämä olisi usein mahdollista, mutta jos Goddardin (1998a) ja Wierzbickan (1991) peräänkuuluttama täydellinen *substitutability* todella olisi huipussaan, eksplikaatioiden tulisi olla millintarkkoja. Tämä ajatus herättää ongelmia niin kieliyhteisön käsitteen (ks. 2.3.1, kysymys 3), semantiikan ja pragmaatiikan erojen (ks. 2.3.1, kysymys 2) kuin myös primitiivien mahdollisen monitulkintaisuuden takia. Eksplikaation tulisi siis olla *exhaustive*; tyhjentävä. Mutta onko mielekäästä olettaa niiden saavuttavan tällaista tasoa?

Maria Koptjevskaja-Tamm ja Inger Ahlgren huomioivat artikkelissaan *NSM: Theoretical, Methodological and Applicational Problems* (2004) NSM:n hyödyn erityisesti silloin, kun halutaan tarkastella kahden lähimerkityksisen sanan eroavaisuuksia. He viittaavat toistuvasti Uwe Durstin artikkeliin *Why Germans don't feel 'anger'* vuodelta 2001. He tarkastelevat esimerkiksi arernten kielen sanan *akiwarre* eksplikaatiota, ja vertaavat sitä jonkin tuntemansa kielen näennäisesti vastaavaan sanaan, esimerkiksi englannin sanaan *anger*.

Oman työni kannalta on oleellista myös se, miten Koptjevskaja-Tamm ja Ahlgren kuvaavat eksplikaatioiden hyötyä toisen kielen puhujalle: Heidän mielestään esimerkiksi sanan *kädet* eksplikaatio tuskin on järkevä hyödyllinen kielenoppijalle, jos tarjolla on myös sanakirjakäännös. Huomio on mielekäs, mutta huomautan, että *kädet* tai *käsi* on sana, joka esiintyy jokseenkin samankaltaisena kielestä toiseen. Voidaan kuitenkin huomata, että sana *käsi* tarkoittaa suomen kielessä entiteettiä, joka voidaan englanniksi kääntää sekä sanaksi *hand* että sanaksi *arm*. Tällöin tarkempi selittäminen on jo toivottavaa.

Oma kysymyksensä on se, ovatko NSM:n puitteissa esitetyt primitiivit todella primitiivisiä. Jatkan aiheesta lisää alaluvussa, jossa käsittelen SLH:ta. Yksittäisen eksplikaation osoittaminen toimimattomaksi, ei-korvattavaksi tai muuten epätarkaksi voidaan tehdä antamalla esimerkkejä, joissa se ei toimi. Näin eksplikaatiota voidaan esimerkiksi tarkentaa tai parantaa, sitä voidaan rajata, mutta yhtäläisesti esimerkkien avulla on mahdollista kritisoida sitä, onko itse eksplikaation muodostaminen reduktiivista parafrasaa käyttämällä ylipäätään mahdollista. Esimerkiksi Geeraerts (2009) esittää kattavia huomioita eksplikaatioista, joiden hän katsoo kaipaavan tarkennusta. On kuitenkin myös syytä muistaa, että kuten Goddard artikkelissaan *Bad Arguments Against Semantic Primitives* (1998b) toteaa, yksittäisen parafrasin epätarkkuus ei missään nimessä tarkoita, etteikö koko metodi toimisi.

2.3.3 Strong Lexicalization Hypothesis ja primitiivien olemassaolo

Alkuun on hyvä käsitellä sitä, miten määritellään alkusana. Sille, että jokin sana on todella NSM:n hyväksymä alkusana ja semanttinen primitiivi, on muodostettavissa kaksi kriteeriä (Goddard 1998a: 57-60): Ensinnäkin sanan tulee löytyä kaikista maailman kielistä. Toisekseen sanaa ei tule voida hajottaa enää pienempiin osiin, eli sille ei pidä voida muodostaa eksplikaatiota. Erityisesti toisen kriteerin merkitys korostuu kehämääritelmiä vältettäessä: Jos sanaa ei voi enää määritellä muin sanoin, ei ole pelkoa siitä, että sen määritelmästä muodostuisi ees taas poukkoileva selitys.

Ensimmäisen kriteerin testaaminen on sikäli hankalaa, että sen tulee tapahtua yksinkertaisesti testaamalla hypoteesia eri tilanteissa tai kielissä. Tämä on työtä, jota NSM:n piirissä (ks. esim. Goddard 1998a) on tehty jo monesta kielestä, myös suomesta. Tavallaan siis tutkielmassani myös testaan tätä hypoteesia ja kriteeriä esittäessäni myöhemmin sellaisia kohtia, jotka herättävät kysymyksiä.

Sen sijaan toisen kriteerin testaaminen on tavallaan yksinkertaisempaa. Goddard (1998a) esittää, kuten aikaisemmin mainittua, että alkusanan primitiivisyyttä voi testata vain määritelemällä se – tämä kutsuu luonnollisesti koettelemaan tätä määritelmää. Lähestyn myös tätä aihetta myöhemmin tutkielmassani, lähteenäni erityisesti Vanhatalon, Tissarin ja Idströmim (2014) artikkeli alkusanojen suomenkielisistä käännöksistä.

Alkusanojen kriteerit eivät kuitenkaan ole täysin vailla ongelmia. Semanttisten primitiivien todeksi osoittaminen on ongelmallista esimerkiksi siitä syystä, että kuten Koptjevskaja-Tamm ja Ahlgren (2004) esittävät, eksplikaation esittäminen jo olemassa olevalle alkusanalle on teorian sisäistä todis-

tamista. Toisin sanoen se sisältää ristiriidan: Jos pyritään esittämään alkusanoin toteutettava eksplikaatio alkusanalle, oletetaan, että alkusanat todella ovat primitiivisiä ja toisaalta pyritään tällä eksplikaatiolla todistamaan tämä oletus vääräksi. Tällaisessa tilanteessa siis oletetaan kahta toisensa poissulkevaa asiaa samaan aikaan. Koptjevskaja-Tamm ja Ahlgren (ibid.) peräänkuuluttavatkin jonkinlaista ulkoista todistamiskeinoa, esimerkiksi Blumczynskin (2013) sanoin *tertium comparationista*, tai Riemerin (2006) kaipaamaa mittakeppiä.

Alkusanojen muodostamiskriteerien ensimmäiseen osaan puolestaan linkittyy oleellisesti niin kutsuttu SLH. Kuten aiemmin mainittua, Goddard itse esittää teoksessaan *Semantic and Lexical Universals: Theory and empirical findings* (1994) NSM:n seitsemännen periaatteen kuuluvan: ”Strong Lexicalization Hypothesis. Every semantically primitive meaning can be expressed through a distinct word, morpheme or a fixed phrase in every language”.

Goddard (ibid.) kuitenkin painottaa, että tämä periaate ei tarkoita jokaisen primitiivin esiintyvän yhtenä ainoana uniikkina muotona jokaisessa kielessä. Päinvastoin, hän painottaa, että nämä primitiivit voivat esiintyä kielessä kompleksisina muotoina, ja ne voivat sisältää elementtejä, jotka toimivat morfeemeina muissa muodoissa. Semanttinen primitiivi siis esiintyy SLH:n mukaan kielessä aina erikseen segmentoitavana muotona, mutta tämä muoto voi itse olla moninainen.

Goddard (ibid.) kuitenkin huomioi, että tämä on hänen mielestään vain vahvin ja todennäköisin vaihtoehto: Hän jättää portin auki sanoen, että primitiivisen merkityksen on mahdollista ilmentyä kielessä myös ei-segmentoitavana osana, kuten reduplikaationa tai ablaut-foneemina.

Nick Riemer (2006) toteaa, että NSM:llä on suurempi tarve kuin monella muulla semanttisella teorialla ottaa huomioon erilaiset merkitykset yhden sanan sisällä. Hän perustelee tätä sillä, että monet kielet vaikuttaisivat olevan ristiriidassa SLH:n kanssa: Riemer nimeää esimerkeiksi japanin kielen konjunktion *-ba*, jonka hän esittää sisältävän kaksi eri primitiiviä: IF ja WHEN. Toinen Riemerin esimerkki on pitjantjataran sana *kulini*, joka on Riemerin mukaan yhtenevä sekä primitiivin THINK että primitiivin HEAR kanssa. Näiden esimerkkien jatkoksi voin esittää myös suomen kielen sanaa *tuntea*, joka niin ikään vastaa kahta NSM:n semanttista primitiiviä: KNOW ja FEEL.

Riemer esittää NSM:n vastauksen näihin huomioihin olleen se, että japanin konjunktio *-ba* ja pitjantjataran sana *kulini* analysoidaan pikemminkin polyseemisinä eri lekseimin ilmentyminä, eikä varsinaisesti kaksi eri muotoa yhdistävinä kokonaisuuksina. Artikkelissaan *Bad Arguments Against Semantic Primitives* (1998b) Goddard tosiaankin painottaa, että kun sanaa *kulini* tarkastellaan lähemmin, voidaan huomata, että sen eri merkitykset poikkeavat huomattavasti sanan syntaktisen käyttäytymisen mukaan. Tämän vastauksen voitaneen olettaa osittain koskevan myös

suomen kielen sanaa *tuntea*. Jos kyse on toisen ihmisen tuntemisesta, englannin kielen merkityksessä *know*, verbin *tuntea* rektio on akkusatiivi: *Minä tunnen sinut*. Sen sijaan, jos puhutaan jonkun tunteen tuntemisesta, englannin sanasta *feel*, partitiivi on yleinen: *Minä tunnen rakkautta*. Huomautan kuitenkin, että tämä jako ei ole yksiselitteinen. Miten suhtautua tämän perusteella lauseeseen ”Minä tunnen piston, kun näen sen”? Tai kun poika kopeloi taskuaan, ja sanoo ääneen: ”Minä tunnen kolikon!”? Tai vielä kolmanneksi esimerkiksi ”Minä tunnen Proustia”.

Riemer jatkaa aiheesta peräänkuuluttaen NSM:ltä vahvaa kantaa polysemiaan. Hän lainaa Goddardin ja Wierzbickan omia sanoja (1994): ”lexicographers agree on at least one mechanical diagnostic of polysemy, namely the possession of mutually exclusive syntactic frames or combinatorial possibilities.” Riemer siis painottaa, että polysemiaa ei voi ottaa tällaisiin kysymyksiin vastaukseksi *ad hoc*, vaan jos polysemia on se tekijä, joka puolustaa SLH:n kantaa, täytyy sen diagnostiikka selvittää täysin.

Yhtä kaikki, tähän periaatteeseen eriävästä syntaktisesta käyttäytymisestä Riemer (ibid.) viittaa termillä *Syntactic Evidence for Polysemy Principle*, eli SEP. Samaan hengenvetoon hän kuitenkin huomauttaa esittäneensä myös SEP:n vastaisia argumentteja, joissa selkeästi saman merkityksen omaavat sanat esiintyvät huomattavan erilaisissa syntaktisissa ympäristöissä.

Riemer (ibid.) painottaakin, että jos SEP on ollakseen toimiva väline polysemian toteamiseen, sen toimiminen ei saisi riippua kontekstista: Jos on olemassa identtisiä sanoja, jotka esiintyvät eri syntaktisissa yhteyksissä, mutta jotka eivät kuitenkaan ole polyseemejä, SEP toimii valikoidusti, ja sen käyttöä tulee täydentää erilaisilla semanttisilla välineillä.

Riemer (ibid.) esittää SEP:n käytön NSM:n kehyksissä olevan hyvin hankalaa myös semanttisten primitiivien osalta. Hän esittää esimerkkejä, joissa englannin kielen sana, NSM-primitiivi *KNOW* esiintyy viidessä erilaisessa syntaktisessa ympäristössä ja niiden perusteella huomioonkin, että nämä syntaktiset ero vaikuttavat olevan tarpeeksi suuret, jotta SEP:tä käytettäisiin erottamaan *KNOW* alkusanan ja primitiivin polyseemiset ominaisuudet ja näin sotimaan itse NSM:ää vastaan.

Lopuksi Riemer (ibid.) vertaa NSM:ää muihin semanttisiin teorioihin yhden oleellisen piirteen osalta. Goddard (2002) esittää, että: ”...taken as a whole, the metalanguage of semantic primes is intended to enable reductive paraphrase of the entire vocabulary and grammar of the language at large, i.e. it is intended to be comprehensive”, ja tähän Riemer tarttuu huomioimalla, että monet muut semanttiset teorat eivät tähtää olemaan kaikenkattavia, kuten NSM Goddardin sitaatin mukaan. Tämä on ongelma siitä syystä, että samassa artikkelissa, jossa Goddard esitti edellisen huomionsa, hän aprikoi myös: ”Perhaps the venture will work out well in some respects and not so well in others; there is no reason to assume a priori that it is an all or nothing affair” (Goddard,

2002). Riemer kritisoi tätä ristiriitaista lausuntoa sanoen, että jos NSM tosiaan tähtää olemaan *The Alphabet of Human Thought*, sillä ei ole varaa tehdä tällaisia myönnytyksiä.

Artikkelissaan *NSM: Theoretical, methodological and applicational problems* (2004) Maria Koptjevskaja-Tamm ja Inger Ahlgren tarkastelevat myös jakoa semanttisiin, tai konseptuaalisiin, universaaleihin sekä leksikaalisiin universaaleihin. Heidän mukaansa esimerkiksi Evans (1994) esittää todisteiden valossa olevan niin, että kayardildin kielessä kaikki NSM:n primitiivit esiintyisivät kyllä semanttisina, mutteivät varsinaisesti leksikaalisina. Esimerkiksi englannin kielen sana ja primitiivi DO saisi kayardildin kielessä eri muodot riippuen siitä, missä yhteydessä se esiintyy. Sellaisenaan sanalla ei kuitenkaan ole kielessä leksikaalista vastinetta.

Ero on sikäli sama kuin esimerkiksi Esa Itkonen (2008) esittämä, että hän katsoo merkityksen jakautuvan seuraavasti: Ensin merkitys hajoaa semanttiseen ja pragmaattiseen merkitykseen (kuten luvussa 2.3.1), jonka jälkeen semanttinen merkitys voidaan vielä hajottaa leksikaaliseen ja grammatikaaliseen merkitykseen.

Itkonen tarkastelee universaalien piirteiden jakautumista eri kantilta kuin NSM: Teoksensa *Maailman kielten erilaisuus ja samuus osa 1* (2008) luvussa VII hän esimerkiksi huomioi grammatikaalisten piirteiden olevan lähtökohtaisesti todennäköisempiä universaaleja kuin leksikaalisten, sillä ne kuvaavat lauseen sisäistä syntaktista suhdetta, ja näin ollen myös niitä elementtejä, jotka muodostavat lauseen merkityksen ja merkityksen ihmisen mielessä. Toisin sanoen Itkonen huomioista on pääteltävissä ajatus, että nimenomaan grammatikaaliset elementit edustaisivat ajattelua kuvaavia universaaleja partikkeleita, käsitettä *alphabet of thought*.

Itkonen (ibid.) esittää myös termin *kvasi-universaali*. Tämä tarkoittaa sellaista elementtiä, jonka yksi kieli koodaa selkeän eksplisiittiseen muotoon, mutta johon toisella kielellä on pääsy vain kiertoilmauksen tai tilannekohtaisen tiedon kautta. Tällainen ilmaus on siis kielennettävissä molemmilla kielillä, mutta toisella sen kielentäminen täytyy tehdä huomattavan monimutkaisesti toiseen verrattuna. Olisi mielenkiintoista tarkastella, missä menee raja universaalien ja kvasi-universaalien välillä, jos sitä peilataan SLH:n sisältämään ajatukseen selkeästä leksikaalisesta koodaamisesta.

Haluan kiinnittää vielä huomioita erilaisiin lähestymistapoihin, jotka ovat jonkinlaisessa ristiriidassa NSM:ssä esitettyjen primitiivien kanssa. Esimerkiksi kielitypologiassa (Anneli Pajunen: luennot Tampereen yliopistossa keväällä 2018, Itkonen 2008) on kyseenalaistettu joitain universaalien luonteeseen liittyviä seikkoja: Pajunen (ibid.) esittää substantiivien ja verbien olevan sanaluokkana kielissä universaaleja, mutta niiden ulkopuolella olevat sanaluokat ovat eri asia. NSM:n alkusanat

kuitenkin sisältävät sanoja jotakuinkin kaikista sanaluokista, esimerkiksi adjektiiveista. Jos on niin, että PIENI koodataan jossain kielessä ei-eksplisiittisenä, vaikkapa verbin kautta, voidaanko sitä silti tarkastella universaalina tai kvasi-universaalina, vai onko se tällöin poikkeus säännössä?

Lisäksi jo mainittu jako leksikaalisiin ja grammatikaalisiin sanoihin on sikäläkin kiinnostava, että koska NSM:n alkusanat sisältävät molempia ja kielitypologia lähestyy grammatikaalisia sanoja huomattavasti todennäköisempinä universaaleina (Itkonen 2008) kuin leksikaalisia, NSM:n alkusanojen kokonaisuudessa vaikuttaa vallitsevan hienoinen epätasapaino, ainakin kielitypologisen tutkimuksen osalta.

Näin ollen painotan itse NSM:n alkusanojen kategorioimisen tärkeyttä ja sitä, että niiden muodostamisperiaatteisiin, jakautumiseen eri kielissä, kielentämiseen ja mahdolliseen hajottamiseen eksplikaatioiden kautta tulisi kiinnittää aina huomiota. Mainittakoon kuitenkin, että Goddard artikkelissaan *Bad Arguments Against Semantic Primitives* (1998) sekä kirjassaan *Semantic Analysis: A Practical Introduction* (1998) tekee eron sanojen erilaisten funktioiden välille.

Eri kielten tutkimusta NSM:n näkökulmasta puolustan eritoten Itkosen (ibid.) huomiolla generatiivisen tutkimuksen epäsojivuudesta lingvistiseen tarkasteluun. Hän siteeraa Bloomfieldiä (1933): ”The only useful generalizations about language are inductive generalizations. Features which we think ought to be universal may be absent from the very next language that becomes accessible.”

Tämä mielestäni perustuu erityisesti siihen, että toisin kuin luonnontieteissä, ihmistieteissä muodostetut säännöt eivät ole aina loogisesti aukottomia. Niitä ei voi eikä pidä tarkastella tavoitteena *quod erat demonstrandum*, vaan inhimillisen muovautuvuuden ja vaihtelun vuoksi sääntöä tulee aina voida täydentää poikkeuksilla, ja sitä täytyy mahdollisesti voida myös muuttaa. Tämä pätee myös kielitieteeseen. Näin, kuten aiemmin mainitsemassani iteroinnissakin, tulokset tarkentavat alkuarvausta ja muokkaavat tietoisuutta sekä ymmärrystä itse asiasta. Kritiikin myötä muovautuva teoria siis vain tekee oman osuutensa hermeneuttisessa kehässä, ja tämän jälkeen sen kuuluu muovautua. Tästä syystä absoluuttisuuden ajatus tuntuu jopa nurkkapatrioottiselta, oli sen noudattajana sitten NSM:ää puoltava tai sitä vastustava teoria.

3. SECOND LANGUAGE ACQUISITION

Second Language Acquisition, eli lyhyemmin SLA, on tutkimusala, joka tähtää selvittämään toisen kielen oppimiseen liittyviä seikkoja. Se on kokonaisuutena varsin laaja, joten tässä tutkielmassa keskityn kuvaamaan sitä, miten se liittyy erityisesti NSM:n alkusanojen käyttöön opetuksessa. Näin tehdäkseeni minun tulee siis erityisesti keskittyä kahteen SLA-tutkimuksen osa-alueeseen: Syötteen eli *inputin* tutkimukseen sekä universaalien käyttöön SLA-tutkimuksessa ja toisen kielen opetuksessa. Syötteellä tarkoitetaan sitä aineista, jota oppijat saavat kohdekielestään erinäisistä lähteistä ja universaaleilla sellaisia piirteitä, joiden voidaan olettaa toistuvan kaikissa maailman kielissä. Perustan tämän luvun erityisesti Rod Ellisin teokseen *The Study of Second Language Acquisition* (1994) sekä Rosamond Mitchellin ja Florence Mylesin teokseen *Second Language Learning Theories* (1998).

Alkuun on hyvä tehdä selväksi joitakuista linjoja. Ensinnäkin Ellis (1994) käyttää termiä *second language*, mutta hän painottaa, että hänen teoksensa huomioita ei ole syytä kohdistaa vain nimenomaan *toisen* kielen oppimiseen: Selvää lienee, että kolmas ja neljäskin kieli noudattelevat samanlaisia kaavoja, mutta ajoittain tämä pätee myös ensimmäiseen kieleen, toisin sanoen lapsen kielenoppimiseen. Tätä huomiota painottavat myös Mitchell ja Myles (1998: 61). Tosin, aikuisena opiskeltava toinen kieli ole lainalaisuuksineen tietenkään täysin verrattavissa ensimmäiseen kieleen. Em. mainitsevat ainakin kolme suurta eroa: **1)** L2-oppijat ovat usein kognitiivisesti kehittyneitä, **2)** L2-oppijoilla on jo tietoutta ja taitoa ainakin yhdestä kielestä ja **3)** L2-oppijoiden motivaatio vaihtelee suurestikin, mutta se harvemmin on samanlaista tarvetta ihmismäiseen kommunikaatioon kuin L1-oppijoilla, ei ainakaan samassa mittakaavassa.

Toinen termin *second language* määrittelyyn tehtävä jako on ero *toisen* ja *vieraan* kielen välillä. Karkeasti tämä ero on seuraava: *Toinen* kieli tarkoittaa kieltä, jota opiskellaan sellaisessa kieliympäristössä, jossa oppijalle on tarjolla kyseisen kielen monipuolista syötettä ja kielenkäyttötilanteita kyseisellä kielellä, esimerkiksi kyseissä maassa. Tämä siis pätee Suomessa asuviin S2-oppijoihin. *Vieras* kieli puolestaan tarkoittaa kieltä, jota opiskellaan tilanteessa, jossa pääsyä samanlaiseen kattavaan syötteeseen ja kielenkäyttötilanteisiin ei ole. Tällaista oppimista voi olla esimerkiksi vieraan kielen opiskelu yliopistossa.

Ellis (ibid.,) jäsentää SLA-kenttää muutamalla karkealla jaolla, jotka on hyvä ottaa huomioon tätä kirjoitettaessa. Ensimmäinen jako on *instructed* vs. *naturalistic*.

Naturalistic eli *luonnollinen* oppiminen tarkoittaa Ellisin mukaan sellaista oppimista, joka tapahtuu ihmistenvälisessä kanssakäymisessä ilman erityistä oppimistarkoitusta. Tällainen oppiminen on siis luonnollista juurikin *toisen kielen* oppimiselle.

Instructed eli *ohjattu* oppiminen puolestaan tähtää sellaiseen oppimiseen, jossa hyödynnetään materiaaleja, ohjausta ja suunnitelmallisuutta. Tätä voi tapahtua sekä *toisen* että *vieraan* kielen oppimisen yhteydessä. Oma tutkimukseni sijoittuu tämän kahtiajaon tälle puolelle, sillä tutkin nimenomaan oppikirjallisuudessa hyödynnettävää NSM:ää, en luonnollista spontaania ihmistenvälistä kanssakäymistä.

Samankaltaista mallinnosta noudattelee jako termeihin *acquisition* ja *learning*. Tästä on kirjoittanut esimerkiksi Krashen (1981).

Seuraava kahtiajako Ellisillä (ibid.) on jako kahteen erilaiseen kielelliseen taitoon: *competence* ja *performance*.

Competence eli *kompetenssi* on Ellisin mukaan lingvistisenä käsitteenä peräisin Chomskyltä (1965). Sillä tarkoitetaan oppijan kykyä käyttää kieltä puhuja-kuulija-tilanteessa ja hänen ymmärrystään kielensisäisistä säännöistä, joita tällaisessa tilanteessa tarvitaan.

Performance eli *performanssi* puolestaan tarkoittaa sellaista toimintaa, jossa kielenoppijan on tarkoitus tuottaa kieltä ilman edellisen kohdan mahdollisuutta sosiaaliseen vuorovaikutukseen.

Omalla tavallaan molemmat liittyvät omaan työhöni: Jos tarkoituksena on pohtia sellaisen mallin mahdollisuuksia, jossa kielenoppija hyödyntää jo tuntemaansa sanastoa ymmärtääkseen ja tuottaakseen monimutkaisempia merkityssisältöjä, kyse voi olla joko performanssista tai vuorovaikutustilanteesta, jossa oppija saa koko ajan palautetta omasta kielenkäytöstään sekä vinkkejä, jotka voivat ohjata häntä oikeaan suuntaan. Tällaisesta toiminnasta on useampia esimerkkejä, mutta esimerkiksi Youtube-kanavan *Abroad in Japan* pitäjä Chris Broad videossaan *Speaking Japanese Fluently in Six Months* (<https://www.youtube.com/watch?v=IOZbJ7PCPk> , aika: 7:52, katsottu 15.3.2018) havainnollistaa tätä käytännössä.

Ellis (1994) esittää SLA-tutkimukselle neljä yleistä kysymystä, jotka asettavat ne puitteet, joissa SLA-tutkimusta tehdään. Nämä kysymykset ovat:

1. What do second language learners acquire?
2. How do learners acquire a second language?

3. What differences are there in the way in which individual learners acquire a second language?

4 What effect does instruction have on second language learning?

Näiden kaikkien käsitteleminen tässä yhteydessä olisi turhaa ja tilaa vievää, mutta erityisesti Ellisin (ibid.: 15) huomiot toisesta kysymyksestä sisältävät työni kannalta kiintoisia elementtejä. Ellis kirjoittaa sellaisista menetelmistä, joita kielenoppijalla on käytössään, kun hän kohtaa ongelmia tai hankaluuksia kommunikoidessaan kohdekielellä. Näitä prosesseja Ellis nimittää laajemmin termillä *communication strategies*. Yksi tällainen strategia on nimenomaan parafraasi: Kohdatessaan tilanteen, jossa kielenoppija ei kykene ilmaisemaan haluamaansa asiaa tarkasti, ja jossa hänellä ei ole esimerkiksi sanakirjaa mukana, hän voi kuvailla tämän asian kiertoilmauksella, *parafrasilla*, jolloin tämä parafraasi toimii kyseisen sanan vaihtoehtona keskustelussa tai esitystilanteessa. Kyseessä on siis tismalleen samanlainen käytäntö kuin aiemmin viittaamassani youtube-videossa.

Viimeinen kahtiajako, jonka Ellisiltä otan, on jako kahteen erilaiseen oppijanäkökulmaan: *external* ja *internal*. Näistä *internal* käsittää sellaisia asioita kuin oppijan L1, hänen henkilökohtaiset kykynsä oppia kieltä, käsitellä esimerkiksi kirjoitettua tekstiä ja niin edelleen. *External* puolestaan koskee niitä oppimisen osa-alueita, jotka ovat säädeltävissä ulkoisesti. Tällaisia osa-alueita ovat esimerkiksi syöte, opetustilanne ja erilaisten opetusmenetelmien hyödyntäminen. Tämä tutkielmani keskittyy vahvasti juurikin *external*-puoleen, sillä se käsittää syötteen mukana myös oppikirjat. Itse mahdollinen NSM-alkusanojen hyödyntäminen käytännössä on puolestaan oppijan taito, joten se sijoittuu pikemminkin *internal*-puolelle.

3.1 SLA:n näkemys universaaleihin

Kuten olen jo maininnut, universaaleilla tarkoitetaan sellaisia piirteitä, joiden voidaan perustellusti olettaa esiintyvän kaikissa maailman kielissä. Joidenkin näkemysten, kuten esimerkiksi NSM:n mukaan tämän on ajatuksena absoluuttinen, toiset omaksuvat pikemminkin näkemyksen, jonka mukaan kyseessä on harmaan eri sävyistä: Piirre, jolla on edustus edes jotakuinkin kaikissa maailman kielissä, on melko universaali.

Ensiksi mainitsen, että SLA:n alaista tutkimusta universaaleista on tehty hieman eri lähtökohdista kuin oma tutkielmani. Yleisesti NSM koskee nimenomaan semanttisia universaaleja,

kun tutkimusta on usein tehty pikemminkin esimerkiksi syntaktisista, kieliopillisista ja foneettis-morfologisista lähtökohdista. SLA:n piirissä tehdyn universaalien tutkimuksen yhteensovittaminen ei siis välttämättä ole ongelmatonta omaan työhöni, mutta pyrin seuraavassa selvittämään, millaisia lainalaisuuksia universaalien tutkimuksesta on löydetty suhteessa SLA-tutkimukseen. Tämä asia kuvastaa myös sitä tosiasiaa, että semanttisia universaaleja ei ole juuri hyödynnetty toisen kielen opetustarkoituksessa. Mahdollisuuksien puitteissa pyrin tarkastelemaan, miten tätä tietoutta universaaleista ylipäättään voisi käyttää hyväksi omassa tutkimuksessani.

Samaan tapaan kuin Itkonen (2008), Ellis esittää hienoisen skeptisyytensä universaalien todellista absoluuttisuutta kohtaan (1994: 417). Jo tässä vaiheessa siis huomataan eroavaisuus NSM:n edustamaan jokseenkin ehdottomaan suhtautumiseen. Ellis esimerkiksi aprikoi, kuinka paljon poikkeavuutta väitetyn universaalien universaaliudessa voidaan sietää: Mikä on se piste, jolloin universaali osoittautuu ei-universaaliksi, eli kuinka monta ko. universaalien olemusta vastustavaa tapaa sallitaan?

Yleisesti Ellis (ibid.: 33) katsoo universaalien hyödyntämisen esittäytyvän eräänlaisena *kielen omaksumislaitteena* (*language acquisition device*), jota oppijat voivat hyödyntää omaksuessaan kohdekielen lainalaisuuksia Ellis (ibid.) lähtee liikkeelle määrittelemällä kaksi eri koulukuntaa, mitä tulee universaalien hyödyntämiseen SLA-tutkimuksessa.

Ensimmäinen joukko ovat niin kutsutut *typologiset universaalit*, joita olen jo sivunnut NSM:n osakseen saamaa kritiikkiä käsittelevässä luvussa 2.3. Typologisilla universaaleilla kuitenkin siis tarkoitetaan sellaisia lainalaisuuksia, joita on saatu selvitettyä vertaamalla useita eri kieli eri kielikunnista.

Toinen joukko on niin kutsuttu *universaali kielioppi*, jota Ellis (1994) kuvaa chomskylaisena koulukuntana. Tämän tutkimuksen puitteissa on keskitytty yksittäisiin kieliin ja pyritty määrittämään sellaisia lainalaisuuksia, jotka puolestaan luovat kieleen kieliopilliset säännöt.

Tässä luvussa käsitelen juurikin *typologisia universaaleja*, sillä katson, että niiden tutkimuksessa virinneitä lainalaisuuksia on mielekkäämpää tarkastella NSM:n valossa kuin *universaalien kieliopin* puitteissa saatuja tuloksia.

Ellis (ibid.) nimeää yhdeksi universaaleihin suhteessa olevan SLA-tutkimuksen suurimmaksi kysymykseksi nk. *omaksumisen loogisen ongelman* (*the logical problem of acquisition*) (ks. myös White 1989). Tämän ajatuksen mukaan se syöte, jolle oppija altistuu, ei ole riittävä, jotta oppija todella voisi omaksua kohdekielen säännöstön. Toisin sanoen tämä syöte ei tarjoa kaikkea informaatiota,

jotta oppija voisi saavuttaa vankan tietoisuuden kohdekielestä. Tästä seuraa, että oppijan on nojauttava muuhun informaation. Tämä informaatio on Ellisin mukaan tietoutta kielellisistä universaaleista: Oppija osaa hyödyntää tietoa omasta kielestään ymmärtääkseen uuden kielen sääntöjä.

Universaalien käyttö opetuksessa ei perustu niinkään suoraan mihinkään erityiseen menetelmään, vaan pikemminkin siihen suhteeseen, mikä universaalien ja oppijat taitojen tai kielitietoisuuden välillä on. Universaaleilla siis selitetään omalta osaltaan sitä, miten oppija todella onnistuu omaksumaan toisen kielen.

Universaalien käyttöön SLA-tutkimuksessa Ellis (ibid.) esittää kolme alaa määrittävää kysymystä:

- ”What influence do typological universals have on the order of acquisition of grammatical features?
- What effect does *markedness* have on learning difficulty?
- How does the typological status of grammatical features in the native and target languages affect L1 transfer?”

Vaikka nimellisesti termi *markedness* esiintyy vain yhdessä näistä kysymyksistä, se määrittää niitä kaikkia. Käytän termistä sen englanninkielistä asua, sillä suoraan suomeksi käännettynä se kuulostaa vähintäänkin erikoiselta (esim. *merkittyys*).

Markedness ja sen vastapari *unmarkedness* rakentavat karkean kahtiajaon sellaisiin kielensisäisiin piirteisiin, jotka tuntuvat olevan ominaisia juurikin kyseiselle kielellä ja sellaisille, joiden voidaan olettaa esiintyvän myös muissa kielissä (Ellis 1994: 319).

Marked tarkoittaa juurikin sellaista piirrettä kielessä, jonka katsotaan olevan ominainen juuri tälle kielelle. *Unmarked* puolestaan tarkoittaa sellaista piirrettä, jonka voidaan olettaa esiintyvän myös muissa kielissä.

Jako on tahallaan polarisoitu, sillä kuten Ellis esittää teoksessaan *Understanding Second Language Acquisition* (1987: 194), jakoa ei voida todella jakaa kahtia, vaan joku piirre voi olla enemmän *marked* kuin toinen. Tämä voi olla myös suhteellista kielten välillä tai ominaista kieliperheille. Esimerkiksi suomen kielen genetiivisubjekti nesessiivirakenteessa finiittimuotoisen verbin yhteydessä on suhteessa vaikkapa tšekin kielen vastaavaan rakenteeseen *marked*: *Minun* (gen.) *täytyy* (yks. 3. p.) *nousta ylös* – *Já* (nom.) *musím* (yks. 1. p.) *vstávat*.

Markednessiin sekä L1-L2-suhteeseen liitten Ellis (1994: 320) lainaa Hyldenstamia (1984, 43): ”Unmarked categories from the native language are substituted for corresponding

marked categories in the target language... Marked structures are seldom transferred, and if they are transferred, they are much more easily eradicated from the target language.” Tähän perustuen Ellis havainnollistaakin asiaa kahdella hypoteesillä: 1) Oppijat siirtävä *unmarked* rakenteita omasta kielestään kohdekieleen, jos vastaava rakenne kohdekieleessä on *marked* ja 2) oppijat vastustavat *marked* piirteiden siirtoa varsinkin, vastaava piirre kohdekieleessä on *unmarked*.

Tätä täydentää Ellisin (1994) lainaama kolmiportainen sääntö Eckmanilta (1977, 321):

- ”Those areas of the TL (*target language*, oma huomio) that are different from the NL (*native language*, oma huomio) and are relatively more marked than in the NL will be difficult”
- “The degree of difficulty associated with those aspects of the TL that are different and more marked than in the NL correspond to the reactive degree of markedness associated with those aspects”
- “Those areas of the TL that are different from the NL but are not relatively more marked than the NL will not be difficult”

Samalla kannalla on myös Zobl (1983), joka esittää, että oppija tukeutuu helposti oman L1:nsä piirteisiin tilanteessa, jossa L2:n säännöt vaikuttavat epäselviltä tai hankalilta.

Ellis (1994: 426) päättää typologisten universaalien käyttöä SLA-tutkimuksessa koskevan lukunsa muutamiiin johtopäätöksiin. Ensinnäkin hän katsoo, että typologisesti *unmarked* piirteet todella vaikuttavat oppijalle helpommin omaksuttavilta piirteiltä kuin sellaiset piirteet, jotka ovat *marked*. Samaa hengenvetoon hän kuitenkin lisää, että universaalit yksin eivät selitä toisen kielen omaksumista.

Itseasiassa Ellis (ibid.) esittää jopa *unmarked* universaalien helppoa omaksumista vastustavan näkökulman: Bardovi-Harlig (1987) toteaa, että vahva syöte voi paikoin jopa ylittää *marked* tai *unmarked* piirteiden voiman: Jos kielellisesti *marked* piirre on vahvasti läsnä oppijalle tarjottavassa syötteessä, hän voi oppia sen nopeammin kuin vastaavan *unmarked* piirteen, joka on syötteessä läsnä vain heikosti. Toisen kielen omaksumista ei siis voi yksin selittää universaaleilla, muttei myöskään syötteen määrällä. Ellis jättää auki kysymyksen: Kumpi on siis enemmän vaikuttava tekijä – syöte vai *unmarkedness*? Ellis tosin katsoo *unmarked* piirteiden olevan todennäköisesti läsnä syötteessä, joten nämä kaksi asiaa eivät välttämättä riitele toistensa kanssa käytännössä. Myös tämä on mielestäni kannustaa oman työni tekemistä: Jos voidaan todeta, että jokin piirre, kuten tässä tapauksessa semanttiset universaalit, todella ovat sekä *unmarked* että vahvasti läsnä syötteessä, ei liene syytä epäillä niiden erittäin sujuvaa opittavuutta.

Luvun viimeisenä huomiona tarjoan Ellisin (1994: 428) esittämän vaihtoehtoisen näkökulman, joka on Ellisin (ibid.) mukaan monen L2-tutkijan omaksuma: *Markedness*- ja *unmarkedness*-suhteet ovat epäsuorassa suhteessa toisen kielen oppimiseen, ja jotta ymmärtäisimme niiden toimintaperiaatteita, tulee meidän keskittyä sellaisiin tekijöihin, jotka tekevät tietystä piirteestä *markedin* tai *unmarkedin*.

Kuten Ellis (ibid, 429) Lightbowta ja Whiteä (1987) lainaa, tietous universaaleista ja *markendessista* saatta olla yhteydessä tietouteen sellaisista ihmismielen piirteistä, jotka sallivat tällaisten asioiden olemassaolon; tämä tietous voi kertoa meille jotain sellaisista malleista, jotka ovat tyypillisiä ihmismielelle.

3.2 Semanttiset universaalit ja SLA-tutkimus – NMS opetuksessa

Kuten aikaisemmin havainnollistin, kielelliset universaalit ovat olleet läsnä SLA-tutkimuksessa paljonkin, ja niiden suhdetta oppimiseen on tutkittu laajalti. On kuitenkin erikseen mietittävä, missä määrin tämän tutkimuksen piirissä saatuja tuloksia voidaan soveltaa semanttisten universaalien käyttöön. Tämä ei ole missään mielessä ongelmatonta, mutta kuten aikaisemmin kuvailin, NSM:n piirissä tehdyllä tutkimuksella on selkeitä yhtymäkohtia kielitypologiseen tutkimukseen. Ensinnäkin menetelmät, joilla NSM:n semanttisia universaaleja haravoidaan, muistuttavat kielitypologiassa käytettyjä menetelmiä: Toisin kuin kieliopillisten universaalien tutkimuksessa, jossa keskitytään tarkasti yhteen kieleen (ks. esim. Ellis 1994, 1987), kielitypologiassa pyritään etsimään yhteneväisyyksiä mahdollisimman monesta määrästä kieliä samaan tapaan kuin NSM:ssä.

Syy, joka puoltaisi NSM:n piirissä tehdyn tutkimuksen yhteensovittamista kielitypologisiin havaintoihin, olisi ajatus siitä, että Ellisin (1994) mukaan kielitypologiset universaalit esittäytyvät mielen kategorioina, ja NSM:n mukaan semanttiset universaalit kuvastavat *ajatusten aakkos-toa* (*alphabet of thought*). Voidaan siis katsoa, että nämä kaksi tutkimussuuntausta tosiaankin voivat kuvata samaa ilmiötä. Tämä on tosin sikäli vain spekulatioita, että tällainen väite tulisi pyrkiä todentamaan tieteellisin keinoin.

Toinen SLA-tutkimuksessa tehdyn kielitypologisen tutkimuksen ja NSM:n piirissä tehdyn semanttisia universaaleja koskevan tutkimuksen yhteneväisyyttä kuvaava piirre on aiemmin esittelemäni käsite *markedness*. Jos oletamme NSM:n semanttisten universaalien todella olevan universaaleja, niitä voidaan myös määrittää Ellisin (1994) käsitteellä *unmarked*. Tällöin ne siis esiintyivät kielestä toiseen jokseenkin selkeinä rakenteina samaan tapaan, kuin SLH:ssa (ks. luku 2.3.3)

oletetaan. On kenties kyseenalaista, ovatko NSM:n semanttiset universaalit absoluuttisia universaaleja, kuten olen selvittänyt luvussa 2.3, mutta kuten puolestaan Ellis (1994) esittää, absoluuttisuus ei välttämättä ole universaaliuden ehdoton tavoite. Olivat semanttiset universaalit absoluuttisia tai ei, ne esiintyvät NSM:n mukaan (ks. esim. Goddard 1998a) erittäin monessa kielessä sellaisinaan. Mielestäni tämä antaa valtuudet ainakin tarkastella niitä käsitteen *unmarked* valossa.

Jos semanttisen universaalit todella ovat *unmarked*, niin Ellisin (1994) arvio niiden yleisyydestä on syytä asettaa luupin alle: Ellis siis olettaa, että *unmarked* piirteet ovat jokseenkin vahvasti läsnä oppijoille tarjottavassa syötteessä. Tutkimukseni perusteella on ainakin nähtävissä, pitääkö tämä arvio paikkansa semanttisten universaalien ja suomenkielisen oppikirjallisuuden perusteella paikkansa. Toisaalta, jos osoittautuu, että semanttiset universaalit ovat aineistossani heikosti edustettuina, voidaan mahdollisesti pohtia sitä, kuinka *unmarked* ne ovatkaan.

Suuri kysymys SLA-tutkimuksessa tehdyn universaaleja koskevan tutkimuksen osalta kuitenkin on se, miten näiden universaalien olemassaoloa voisi opetuksessa hyödyntää. Tähän mennessä esittämäni huomiot SLA-tutkimuksesta koskevat pikemminkin sitä, kuinka hyvin oppijat tosiaan oppivat universaaleina pidettyjä *unmarked* piirteitä erityisesti *marked* piirteisiin verrattuna. Tämä ei vielä vastaa kysymykseen siitä, miten niitä voisi käyttää.

Ellisin teoksesta *The Study of Second Language Acquisition* (1994) on kuitenkin löydettävissä huomioita, jotka ainakin epäsuorasti viittaavat universaalien olevan myös hyödyllisiä opetuksessa. Erityisesti tällaiseksi huomioksi nostan Ellisin mainitseman *communication strategy* (1994).

SLA-tutkimuksesta tekemäni selvityksen perusteella voidaan myös perustella sitä, miksi universaaleja voisi käyttää kommunikatiivisen strategian kehittämisessä. Kuten Goddard kirjoittaa (2001), NSM-alkusanojen oletetaan sisältyvän kieltä oppivan ihmisen mentaalileksikkoon jo varhain. Jos myös oletetaan, että ne ovat *unmarked*-piirteinä myös läsnä kohdekielellä tarjottavassa syötteessä, tarjottu syöte siis kohtaa kohdekielen syötteen hyvin syvällä tasolla: Semanttiset universaalit, tai niin niiden vastineet oppijan omalla kielellä, kuuluvat jo oppijan ymmärtämään ainekseen, jolloin niiden voidaan olettaa siirtyvän kohdekieleen jokseenkin vaivatta, kuten Eckmanin (1977) hypoteesi olettaa: Niiden piirteiden, jotka ovat kohdekielessä erilaisia, mutta jotka eivät ole mainittavasti enemmän *marked*, voidaan olettaa siirtyvän kohdekieleen jokseenkin vaivatta.

Näiden havaintojen perusteella voin esittää hypoteesin, että oppijat todella voivat oppia semanttisia primitiivejä vastaavat sanat kohdekielellään suhteellisen vaivatta, jolloin pohja niiden käyttämiselle kommunikatiivisia strategioita luotaessa olisi muodostettavissa hyvinkin sujuvasti.

SLA-tutkimusta ja universaaleja koskevan lukuni päätteeksi esitän vielä huomioita siitä, millainen tämä NSM:n semanttisia primitiivejä hyödyntävä kommunikatiivinen strategia olisi.

Kuten Goddard esittää (1998a), NSM:n avulla muodostetut eksplikaatiot muistuttavat mahdollisimman paljon luonnollista kieltä omine sääntöinen ja kieliopillisine piirteineen. Tällöin on hyvä pitää mielessä, että 60-70 sanan lista yksinään ei riitä oppijalle olemaan pohja kielelliselle taidolle. Jotta oppija voisi käyttää strategiaa, jossa hän hyödyntää oppimiaan NSM:n piirissä tarkasteltuja semanttisia primitiivejä, täytyy hänellä olla tietoa myös kohdekielen rakenteesta. Toisin sanoen hänen täytyy osata muodostaa lause annetuista sanoista.

NSM:n alkusanoja hyödyntävä strategia ei siis ole mikään supermetodi kielenoppimiselle: Se on selviytymiskeino, johon oppija voi nojata tilanteissa, joissa hänellä ei ole tarvittavaa sanastoa tai pääsyä esimerkiksi sanakirjaan kyseisellä hetkellä. Se, että oppija hallitsee NSM:n alkusanat kohdekielellä takaa vain ymmärtämisen alimman tason: Oppija voi tai oppijalle voidaan selittää termi tai käsite siten, että se nojaa varmasti ymmärrettyyn ainekseen.

Kuten olen jo luvussa 2.1.2 kuvannut, molekyylit voivat osoittautua hyvinkin tehokkaiksi palasiksi NSM:n avulla muodostetun strategian toteuttamisessa. Ne myös omalta osaltaan vastaavat Piotr Blumczynskin (2013) ja Nick Riemerin (2006) peräänkuuluttamaan ymmärrettävyyteen eksplikaatioissa. Tämän vuoksi niiden ja muidenkin oppijoiden jo ymmärtämien sanojen voidaan tapauskohtaisesti ehdottomasti sallia kuuluvan myös NSM:n perusteella kehitettyyn kommunikatiiviseen strategiaan.

Katson, että tietouden näistä sanoista olisi hyvä piirtyä oppijan ymmärtämään ainekseen mahdollisimman varhain, sillä NSM:n perusteella monimutkaisemman käsitteet loppujen lopuksi rakentuvat niistä. En kuitenkaan oleta, että vaikka alkeisoppikirja sisältäisi kaikki NSM:n alkusanat syötteessään, opiskelija voisi kyseisen kirjan opiskeltuaan käyttää näitä sanoja automaattisen sujuvasti kielessään. Kielenoppiminen, kuten jo puhtaasti SLA-tutkimuksen laajuudesta käy ilmi, on hyvin monipuolinen ja laaja tieteenala, ja sellaiset näkökulmat, jotka näyttävät liian hyvältä ollakseen totta todennäköisesti ovat sitä.

4. AINEISTO

Aineistonani käytän siis suomenkielistä S2-oppikirjallisuutta. Oppikirjojen merkityksestä voidaan käydä keskustelua, eikä ole tietenkään itsestään selvää, että toimiva oppikirja johtaa onnistuneeseen oppimiskokemukseen, mutta perustelut oppikirjojen tutkimukselle ovat mielestäni vankat: Tutkimukseni on ensimmäisiä vastaavia tutkimuksia ainakin suomenkielisessä kontekstissa, ja tästä syystä on mielestäni hyvä pitäytyä selkeän standardoidussa materiaalissa, joka on lisäksi taajaan käytössä. Kuten Tanner (2012) esittää, oppikirjat eivät missään nimessä ole ainoa materiaali, jota opetuksessa käytetään: Monet opettajat kehittävät omaa materiaaliaan jatkuvasti, enkä itsekään ole tähän poikkeus. Lähtökohtaisesti oppikirjat kuitenkin ovat mielestäni selkein aloituspiste.

Opetushallituksen dokumentissa *Kommentoitu luettelo maahanmuuttajataustaisten opetuksen ja koulutuksen materiaaleista* vuodelta 2012 tehdään laajasti selkoa siitä, minkälaista kirjallisuutta on tällä hetkellä käytössä S2-opetuksessa. Koko tämän aiheiston tutkiminen olisi erittäin vaativa tehtävä, joten sitä on tarpeen rajata.

Ensinnäkin tarkennan tutkimukseni koskemaan aikuiskoulutuksessa käytettäviä alkeistason oppikirjoja, joissa ei ole erityisiä painotuksia, kuten esimerkiksi opetushallituksen *Elämä tarinaksi* -teoksessa, joka näkee kielen kertomuksina yksittäisten rakenteiden ja piirteiden sijaan. Jakoa aikuisille tarkoitettuihin kirjoihin perustelen sillä, että kuten luvussa 1.1 kerron, alkusanojen voidaan katsoa ilmaantuvan kieleen jo varhain. Painottamalla aikuisille suunnattuja oppikirjoja pyrin myös mahdollisimman hyvin varmistamaan sen, että näiden kirjojen kohderyhmällä on jo kattava kielellinen kompetenssi ainakin äidinkielellään. Alkeistason oppikirjallisuutta puolestaan perustelen sillä, että NSM:n alkusanat ovat luonteeltaan yksinkertaisia, ja voitaneen olettaa, että suuri osa niistä sisältyy jo varhaisessa vaiheessa opetettavaan ainekseen.

Otan rajauksessani mallia Johanna Tannerin (2012) väitöskirjasta *Rakenne, tilanne ja kohteliaisuus – pyynnöt S2-oppikirjoissa ja autenttisissa keskusteluissa*, jossa aineiston rajausta perustellaan erityisesti sillä, että uusintapainosten määrä korreloi kyseisen kirjan suosittuuden kanssa. Olen rajannut aineistoni koskemaan vain 2000-luvulla tehtyjä tai uusittuja kirjoja. Näillä perusteilla olen valinnut aineistokseni yhden laajalti käytetyn S2-oppikirjan, josta on otettu useita painoksia. Katson aineiston olevan riittävä, sillä vaikka osa työstäni onkin mekaanisesti alkusanojen esiintymien kartoittamista, työssäni on vahvasti mukana laadullinen aspekti. Kuten tutkielmani johdantoluvussa 1 käy ilmi, en pelkästään käsittelen NSM-alkusanojen määrää tässä oppikirjassa, vaan lisäksi pyrin arvioimaan koko NSM-sanastoa suomen kielen ja toisen kielen opetuksen kannalta. Valitsemani kirja on Hyvin menee! 1 – Suomea aikuisille -kirjan 9., uudistettu painos vuodelta 2010. Sen

tekijät ovat Satu Heikkilä ja Pirkko Majakangas. Sen kustannuksesta on vastannut kustannusosakeyhtiö Otava.

4.1 Kirjan esittely

Olen jakanut kirjan *Hyvin menee I* sisällön neljään osa-alueeseen: tekstit, tehtävät, sanastoa opettavat kokonaisuudet ja kielioppikokonaisuudet. Tekstit on kirjassa selkeästi esillä ja merkitty usein vihreällä värillä. Tehtäväkokonaisuudet on ripoteltu kappaleisiin, ja ne on merkitty vihreillä numeroituilla palluroilla. Sanastoa opettavat kokonaisuudet ovat luonteeltaan sellaisia, että ne kiinnittävät huomion johonkin tarkkarajaiseen kieliopilliseen, käytännölliseen tai sanastolliseen kokonaisuuteen. Ne on merkitty kirjassa sinisellä värillä. Kielioppikokonaisuudet on merkitty kirjaan punaisella. Ne sisältävät jokseenkin monimutkaista kieltä ja teknisiä kieliopillisia termejä.

Rajaan kirjan sisältöä siten, että en ota työssäni huomioon kielioppikokonaisuuksia kahdesta syystä. Ensinnäkin ne sisältävät niin monimutkaista kieltä, että en oleta oppijoiden lukevan niitä itsenäisesti ja sen myötä niiden aines on epärelevanttia. Toisekseen kielioppikokonaisuudet sisältävät lukuisia saman alkusanan käyttöyhteyksiä jokseenkin epäluonnollisessa ympäristössä ja toistoa, mikä vääristää asetelmaani.

5. MENETELMÄT

Tutkimuksessani kartoitan siis ensinnäkin kuinka paljon NSM:n sanastosta kirjoissa esiintyy. Havainnoin sitä, missä vaiheissa sanat tuodaan tekstiin, eli kuinka aikaisin mitkäkin alkusanat alkavat sisältyä oppijoille tarjottavaan ainekseen. Havaintojeni pohjalta pohdin sitä, miksi mitkäkin sanat esiintyvät tekstissä niin paljon kuin ne esiintyvät ja mistä syistä ne mahdollisesti tuodaan mukaan syötteeseen silloin, kun ne tuodaan.

Olen muokannut Vanhatalon, Tissarin ja Idströmin (2014) kääntämästä alkusanalistasta uudelleen jaotellun alkusanapaketin jota käytän apunani kartoittaessani aineistoa. Tämän pake-
tin luomisessa olen ottanut mallia myös Tissarin ja Vanhatalon päivitetystä alkusanalistasta (Virittäjä, 2017/2) sekä Seija Tuovilan väitöskirjassa (2005). Itse alkuperäisissä alkusanalistoissa ei ole minusta mitään sellaista, mitä täytyisi niinkään parantaa, vaan kyse on niiden soveltamisesta käytäntöön. Käydessäni läpi oppikirjan tekstisyötettä minun tulee voida tehdä ratkaisuja sen suhteen, mikä sana lasketaan tietyn semanttisen primitiivin esiintymäksi ja mikä ei. Tämä on ajoittain tehtävä, joka vaatii harkintakykyä, sillä jotkin sanan eivät sovi semanttisten primitiivien kategorioihin ongelmitta. Pyrin tässä luvussa esittelemään tähän syitä, sekä esittämään tekemäni ratkaisut mahdollisimman yksityiskohtaisesti. Vaikka olen usein pyrkinyt perustelemaan tekemäni ratkaisut mahdollisimman hyvin, ne ovat viime kädessä omiani, ja hyväksyn sen, että niistä voidaan olla myös eri mieltä.

Virittäjän numeron 2/2017 verkkoversiossa on myös ilmestynyt Tissarin ja Vanhatalon lista alkusanoista niiden kaanon-konteksteissa. Tällä tarkoitetaan sellaisia käyttöympäristöjä, joissa alkusanan voidaan olettaa esiintyvän. Asia ei ole aina yksiselitteinen; joskus polyseemisen sanan voi erottaa alkusanasta vain sen esiintymisympäristö. Olen pyrkinyt huomioimaan Tissarin ja Vanhatalon (ibid.) listan omaa työtäni tehdessäni, vaikka on syytä mainita, että oma lähestymistapani aineistooni on hivenen erilainen: En pyri niinkään tarkastelemaan niitä kielellisiä ympäristöjä, joissa alkusanat esiintyvät – ellei tämä ympäristö ole aivan oleellinen alkusanan polyseemisten tai homonyymisten ominaisuuksien tarkastelun vuoksi – vaan keskityn pikemminkin erittelemään alkusanojen polyseemiset, homonyymiset ja syntaktiset elementit aineistolähtöisesti. Em. kirjoittajien lista alkusanoista niiden tyypillisissä konteksteissa on työssäni liitteenä (liite 1).

Tässä luvussa puhun usein *semanttisista primitiiveistä*. Tällä tarkoitan sellaisia *langue*-tason sanoja, joille kielistä etsitään *parole*-esiintymiä. Semanttinen primitiivi on siis idea, ja alkusanat sekä niiden alloleksit sen ruumiillistumia kielenkäytössä.

Itse tutkimukseni aineiston käsittely koostuu seuraavista vaiheista:

1. Kartoitan NSM-alkusanat oppikirjan tekstikappaleista, tehtävistä ja sanastoa käsittelevistä kokonaisuuksista ja tilastoin ne mahdollisimman tarkasti.
2. Muotoilen kirjasta yleiskuvan ja tarkastelen sitä, milloin mikäkin alkusana tuodaan syötteen mukaan sekä kuinka yleinen se missäkin kirjan osa-alueessa on.
3. Pohdin syitä jakautumisen eroille niin kirjan eri osa-alueiden kuin kirjan kappaleidenkin välillä.
4. Pohdin syitä sille, miksi alkusanat jakautuvat aineisooni niin kuin ne jakautuvat.

5.1 Syyt ratkaisujen taustalla

Tässä luvussa selitän niitä ratkaisujani, joita olen katsoakseni joutunut tekemään NSM:n käytön suhteen. Selitän myös, miksi tällaisia ratkaisuja täytyy tehdä. Kuten aiemmin sanottua, käytän työni pohjana Vanhatalon, Tissarin ja Idströmin (2014) muodostamaa NSM:n alkusanojen suomennosta. Tissari ja Vanhatalo kuitenkin esittelivät *Virittäjän* vuoden 2017 2. numerossa päivitettyjä huomioita NSM:n alkusanaparadigmaan. Lisäksi samassa artikkelissa esiteltiin eräitä valintoja, joita Seija Tuovila oli tehnyt väitöskirjassaan *Kun on tunteet. Suomen kielen tunnesanojen semantiikkaa* vuonna 2005.

Suurin yksittäinen syy sille, että työssäni tulee paikoin tehdä uudenlaisia ratkaisuja NSM:n alkusanoja käytettäessä on se, että työni ei niinkään ole semanttinen, vaan alkuperältään semantiikkaan pohjaavaa tutkimusta soveltamaan pyrkivä tutkielma. Tämä tarkoittaa sitä, että työni tavoite on jokseenkin eri kuin esimerkiksi toimivien semanttisten selitysten luominen. Tämä asia minun tulee ottaa huomioon esimerkiksi sellaisissa tilanteissa, joissa en ole varma, onko jokin sana hyväksyttävä tietyn semanttisen primitiivin edustukseksi aineistossa.

Kategorisesti voin sanoa, että usein olen pyrkinyt pikemminkin sallimaan kuin rajoittamaan synonyymejä tai hyvin lähimerkityksisiä sanoja aineistooni. Syy tähän on käytännöllinen: Jos oppija haluaa kuvata puheessaan jonkin haluamista tai tahtomista, ei ole niin väliä, kumpaa sanaa hän käyttää; viestin menee todennäköisesti perille. Yleinen kahtiajako on esimerkiksi jako abstraktiin tai ei-laskettavaan ja laskettavaan. Tämä havainnollistuu vaikkapa sanoissa *osa* ja *pala*, joista vain *osa* on laskettu alkusanaksi. Näiden käyttäminen toistensa sijaan on virhe, mutta virheenä niin marginaalinen, että se tuskin vaikeuttaa paljoakaan ymmärretyksi tuleamista.

Ennen valintojeni yksityiskohtaista esittelyä kirjoitan hieman niiden taustalla olevista teemoista: alloleksisyydestä, taivutuksesta, homonymiasta ja polysemiasta.

5.1.1 Alloleksisyys ja taivutus

Ensimmäiseksi keskityn suomen kielen monipuoliseen taivutussysteemiin. Teen eron nimitaivutuksessa kieliopillisen ja leksikaalisen sijataivutuksen välillä, sekä komparaation osalta. Lisäksi otan huomioon erot yksikön ja monikon välillä. Verbitaivutuksessa otan huomioon persoonataivutuksen, tempukset ja modukset. Lisäksi selitän tekemiäni ratkaisuja infiniittimuotoisten verbien osalta.

Ensinnäkin on hyvä painottaa eroa kieliopillisten ja leksikaalisten sijojen välillä nimitaivutuksesta puhuttaessa: Kieliopilliset sijat auttavat määrittämään sanan syntagmaattisen roolin. Tällöin ne eivät varsinaisesti sisällä sellaista semanttista ainesta, jonka perusteella ne täytyisi erottaa niistä sanoista, joita tutkimuksessani käsittelen. Pidän niitä siis puhtaina allolekseinä, eli semanttisten primitiivien kanssa yhteneväisinä. Leksikaaliset sijat, ja eritoten paikallissijat ovat hieman pohdintaa vaativa asia.

Paikallissijoilla on merkitys, joka tosiaankin kuvaa sanan paikallisuutta, mutta monella on myös kieliopillinen merkitys, tai ainakin rektion määrittämä vaadittu muoto. Jälkimmäinen tapaus putoaa edelleen samaan kategoriaan kuin kieliopilliset muodot yleensä, eli en näe syytä karsia niitä aineistostani pois. Tämän lisäksi katson myös, että muutkin paikallissijat on laskettava alloleksisyyden piiriin: Perustelen tämän sillä, että jos tekisin työtäni esimerkiksi englanniksi, katsoisin ehdottomasti ilmaisun *from you* todella sisältävän semanttisen primitiivin *you* ilmentymän. Tällöin ilmaisussa oleva *you* siis laskettaisiin aineistooni mukaan. Katsoakseni ei siis ole mitään syytä karsia ilmaisuja *sinulta* tai *sinusta* pois aineistostani. Näin ollen lasken siis kaikki sijataivutustapaukset aineistooni kuuluviksi. On myös hyvä mainita, että NSM:n alkusanojen paradigma sisältää myös jonkin verran taivutusaineksia: Vanhatalo, Tissari ja Idström (2014) katsovat englanninkielisen alkusanan KIND vastineen olevan suomen kielessä mm. pääte -LAINEN/-LÄINEN. Lisäksi Tissari ja Vanhatalo (2017) nimeävät inessiivipäätteen -SSA/-SSÄ englanninkielisen alkusanan WHERE ~ PLACE esiintymäksi.

Ero yksikön ja monikon välillä on näiden edellisten huomioiden jälkeen mielestäni jopa triviaali. Lasken poikkeuksetta yksikkö- ja monikkomuodot kyseisten semanttisten primitiivien edustajiksi. Oikeastaan ainoat tilanteet, joissa tämä asia tarvitsee edes huomioida, ovat englanninkieliset alkusanat PEOPLE ja WORDS. Näistä kuitenkin lisää luvussa 5.2.

Mitä tulee verbitaivutukseen, edellisessä kappaleessa mainitsemani syyt pätevät myös niin persoonataivutukseen kuin moduksiinkin. Esimerkiksi englannin ilmaus *s/he probably is* sisältää primitiivin *to be* esiintymän. Näin ollen sen sisältää myös sekä persoonataivutuksen että modusvaihtelun alainen *hän lienee*.

Infiniittimuotoisten verbien mukaan aineistooni laskeminen perustuu myös useampaan asiaan. Ensinnäkin suomen kielessä sanakirjoissa esiintyvä verbimuoto on verbin A-infinitiivi. Olisi outoa erottaa se siis aineistosta. Toisaalta, esimerkiksi partisiippimuodot ovat usein jopa sellaisia, joita ei vaikkapa englannin kielessä voisi ilmaista: *Tulin tehneeksi kaikenlaista*. Katson, että tämä lause sisältää semanttisen primitiivin *tehdä* ilmentymän, vaikka se onkin translatiivitaivutettu aktiivin 2. partisiippi.

Verbien tempustaivutuksesta olen jo esittänyt luvussa 2.1.2 NSM:ssä vallalla olevan näkemyksen siitä, että esimerkiksi ollakseen kieliopillinen, eksplikaation tulee voida sisältää myös verbien eri aikamuodot. Pidän näin ollen myös tempustaivutusta aineistooni kuuluvana piirteenä. On tosin syytä huomata, että suomen kielessä perfekti ja imperfekti sisältävät myös sanan *olla*, jolloin sen kategorisoimista aineistooni tulee vielä tarkastella lähemmin. Myös tästä lisää luvussa 5.2.

Johtopäätöksenä siis totean, että lasken työni aineistoon jotakuinkin kaiken taivutetun kieliaineksen. Tämä johtuu erityisesti siitä syystä, että jos aion todella tilastollisesti selvittää alkusanojen kerrostumia oppikirjallisuudessa, en voi perustellusti esittää, etteikö *sanottu* olisi alkusanan *sanoa* esiintymä. Kyseisen sanan muoto palvelee lauseessa niin montaa eri tarkoitusta, esimerkiksi syntaktista, etten voi katsoa sen ylipäätään välttämättä esiintyvän samankaltaisessa yhteydessä taivuttamattomana. Olen siis ehdottomasti sitä mieltä, että taivutusta ei voi erottaa suomen kielestä, jolloin katson, että kaikki taivutetut muodot edustavat sellaista alloleksisyyttä, että lasken ne mukaan aineistooni. Ratkaisuani tukee, kuten sanottua, myös se, että pyrin lähestymään aineistoani hyvinkin käytännöllisesti. Liian vahva pedanttisuus ei mielestäni edistä NSM:n käyttöä opetuksessa kovinkaan paljoa. Kokonaisuudessaan alloleksisyyden hyväksymistä perustelen myös sillä syyllä, että jos kyseisen semanttisen primitiivin edustaja todella esiintyy aineistossa, se tarkoittaa sitä, että kyseinen primitiivi on aineistossa läsnä. Tällöin oppija ymmärtää, mistä primitiivistä on kyse. En siis juurikaan erottele aineistossani liki synonyymisiä esiintymiä kuten *haluta* ja *tahtoa*.

5.1.2 Homonymia ja polysemia

Homonymia ja polysemia muodostavat yhdessä ilmiön, joka tulee ottaa huomioon siinä, miten määrittelen aineistooni mukaan otettavan aineksen. Näillä termeillä tarkoitetaan sitä, että sanat voivat

suomen kielessä ja muissakin kielissä olla kirjoitusasultaan tismalleen samanlaisia, vaikka ne tarkoittavat eri asioita. Jako homonymian ja polysemian välillä ei ole aina selkeä, mutta usein se on määriteltävissä. Karkeasti määritellen homonymialla tarkoitetaan esimerkiksi taivutuksesta johtuvaa samanmuotoisuutta (yks. gen. sanasta *syöpä* on *syövän*, joka on myös genetiivimuoto verbin *syödä* VA-partisiipista) ja polysemialla leksikaalista monimerkityksisyyttä (*viini* voi tarkoittaa sekä alkoholipitoista rypäleistä valmistettua juotavaa että jousiampujan selässään kantamaa nuolikoteloä). Eron voi paikoin havaita jopa taivutuksesta (kuten edellisessä esimerkissä *viinin* - *viinen*), mutta se on usein myös selvä kontekstista.

Luvussa 5.2 käsittelen aineistostani löytyneet homonymian tai polysemian tapaukset, joihin minun täytyy kiinnittää erityistä huomiota.

5.2 Käyttämäni alkusanalista

Tässä luvussa esittelen sen alkusanalistan, jota olen käyttänyt aineistoani käsitellessäni. Lista pohjana, kuten sanottua, on Vanhatalon, Tissarin ja Idströmin (2014) ehdottama englanninkielisistä alkusanoista käännetty paketti, mutta olen tehnyt siihen joitain lisähuomioita. Suurin osa listan sanoista on juuri sellaisessa asussa kuin em. ne esittelevät. Huomioni keskittyvät lähinnä semanttisten primitiivien alloleksien määrän lisäämiseen: Kun olen aineistoa käsitellessäni törmännyt sellaiseen sanaan, joka ei varsinaisesti kuulu NSM:n alkusanojen paradigmaan, mutta joka mielestäni joko vastaa tai on hyvin samanmerkityksinen jonkin alkusanan kanssa, olen tapauskohtaisesti tehnyt sen perusteella ratkaisun: Tuleeko tämä sana ottaa mukaan aineistooni, vai pitääkö se jättää siitä pois. Esimerkkinä tällaisesta sanasta olkoot vaikka sanat *väki* ja *joukko*. Semanttinen primitiivi IHMISET kuuluu NSM:n englanninkieliseen paradigmaan, joten mielestäni sanat *väki* ja *joukko* voidaan hyväksyä sen allolekseiksi ainakin työssäni käyttämäni matalan kynnyksen alloleksisyyttä noudattaen.

Olen myös jonkin verran erottanut alun perin yhteen semanttiseen primitiiviin kuuluvia sanoja kokonaan omiksi alkusanoikseen. Tällainen sana on esimerkiksi englanninkielinen IT, jonka suomennos on em. listassa TÄMÄ ~ SE. Olen selittänyt kyseisten sanojen jakoperusteet tarkemmin itse listauksessani, mutta yleisenä ilmiönä on helpompaa jakaa yksi semanttinen primitiivi kahdeksi eri alkusanaksi, kuin yhdistää kaksi alkusanaa yhdeksi. Tämä johtuu siitä, että aineistoa käsitellessä on helppoa laskea kaksi sanaa yhteen kategoriaan kuuluviksi, mutta jotta voi hajottaa yhden sanan kahtia, täytyy kahlata koko aineisto uudestaan läpi.

Yleisiä ongelmia, jotka aineistossani ovat nousseet, ovat jako abstraktiin ja konkreettiseen tilanteissa, joissa vain toinen lasketaan alun perin alkusanaksi (OSA ~ PALA), vastarakenteet, joissa molemmat sanoista voitaisiin esittää negaatiolla toisen kanssa, mutta joista vain toinen kuuluu alkusanojen paradigmaan (SAMA ~ ERI) sekä fraasit ja kiteytymät, jotka oikeastaan vastaavat jonkin tietyn alkusanan olemusta (OLLA MIELTÄ → AJATELLA, *että*). Eritoten viimeksi mainittuihin ilmentymiin teen tietoisesti poikkeaman siitä, mitä Goddard (1998a) kehottaa: Hän sanoo, että jos jokin sana periaatteessa yhdistäisi kaksi alkusanaa, ei niitä molempia voida katsoa tähän sanaan kuuluviksi. Olen asiasta ainakin oman työni puitteissa eri mieltä: Jos oppija törmää vaikkapa ilmaisuun *Minun tekee mieli olutta*, on se hänelle helppo selventää ilmaisuksi *Minä haluan olutta*, vaikka ilmaisuilla tarkempi semanttinen merkitysero onkin. Katson kuitenkin, että oppiessaan tämän ilmaisan, oppija reagoi myös sanoihin *tehdä* ja *mieli*, eli niitä ei voi kokonaan sivuuttaa. Tosin, tämäkään asia ei ole mustavalkoinen: Siinä missä *Minun tekee mieli olutta* katsoakseni tosiaankin sisältää myös sanat *tehdä* ja *mieli*, ei esimerkiksi sanan *lohikäärme* välttämättä sisällä enää ajatusta lohesta ja käärmestä. Tällöin sana on selkeästi enemmän kuin osiensa summa. Kokonaisuudessaan kuitenkin painotan, että käsittelen tällaiset tilanteet tapauskohtaisesti. Tämä omalla tavallaan kvasiholistinen näkemykseni on vain yleinen linjaus: Kokonaisuus voi olla enemmän kuin osiensa summa, mutta osat voivat olla enemmän kuin kokonaisuus.

Olen järjestänyt listan siten, että sijoitan sanat yläkategorioihin. Näitä kategorioita on 9. Ne ovat **1)** ihmiset, oliot ja niihin liittyvät sanat, **2)** muut nominit, **3)** määrää kuvaavat sanat, **4)** adjektiivit, **5)** verbit, **6)** aikaa kuvaavat sanat, **7)** paikkaa kuvaavat sanat, **8)** kieliopilliset sanat sekä **9)** muut taipumattomat sanat. Olen tehnyt jokaisesta alakategoriasta oman alalukunsa. Lukujen loppuun olen lisännyt sulkeisiin sellaisia sanoja, jotka mielestäni voisivat suomen kielessä olla alkusanoja, mutta jotka eivät kuulu NSM:n paradigmaan. En ole laskenut näitä sanoja mukaan aineistooni, ne ovat läsnä lähinnä huomioina. Olen myös tehnyt oman lukunsa kysymyssanoista, joita en kuitenkaan ole laskenut mukaan varsinaiseen aineistooni.

Joidenkin taipumattomien sanojen rinnalla saattaa esiintyä esimerkiksi nomini, koska ne on semanttisen vastakkaisuuden vuoksi järkevää esittää rinnan. Olen alleviivannut sellaiset sanat, jotka olen lisännyt listaan itse. Lista on seuraavanlainen:

5.2.1 Ihminen, olio ja niihin liittyvät sanat

MINÄ & SINÄ

Asia, joka tulee ottaa huomioon, kun tilastoidaan alkusanojen MINÄ ja SINÄ esiintymiä, liittyy siihen, kuulisiko aiemmin esittämieni syiden perusteella laskea mukaan myös sanan *me* ja *te*. Olen päättänyt ratkaisuun, jossa jätän nämä tilastoinnin ulkopuolelle, sillä kyse ei kuitenkaan ole monikkotaivutuksesta, jonka alaiset tapaukset lasken aineistooni mukaan. Primitiivien englanninkieliset vastineet ovat I ja YOU.

JOKU ~ IHMINEN ~ HÄN ~ HENKILO

Kokonaisuudessaan tämän semanttisen primitiivin edustus on mielestäni laaja. Vanhatalo, Tissari ja Idström (2014) ovat päätyneet katsomaan myös suomen kielen yksikön kolmannen persoonan persoonapronominin *hän* lukeutuvan tähän joukkoon, enkä näe syytä kiistellä heidän valintansa kanssa. Alkusanan englanninkielinen vastine on SOMEONE. Oikeastaan kategoria viittaa siis sellaisiin ihmisiin toimijoihin, joihin viitataan kolmannessa persoonassa. Tämä tarkoittaa, että esimerkiksi lause *Minä olen ihminen* ei sisällä tämän semanttisen primitiivin esiintymää, sillä tällöin on kyse ihmisestä lajina.

Olen laskenut mukaan myös aloleksiksi sanan *henkilö*.

IHMISET ~ VÄKI ~ JOUKKO

Tämä on primitiivi, jonka sovittaminen suomen kieleen on mielestäni ongelmallista. Primitiivin englanninkielinen vastine on PEOPLE, joka sisältyy siis monikkomuotoisena NSM:n alkusanojen listaan. Olen tästä lähtökohtaisesti jonkin verran eri mieltä: Alkuperäinen Vanhatalon, Tissarin ja Idströmin (2014) ehdotus primitiivin suomenkieliseksi vastineeksi on IHMISET, joka on vain sanan *ihminen* monikkomuoto. Koska katson monikkomuotojen edustavan samaa semanttisen primitiivin kategoriaa kuin yksikköjenkin, ei tämä sana näyttäydy minulle alkusanana. Myös Vanhatalon mielestä sana on ongelmallinen (sähköpostikeskustelu).

Aiheesta voidaan käydä yhä keskustelua. Olen ottanut sanan kuitenkin aineistooni mukaan; aineistoani voi halutessaan käsitellä myös niin, että jättää tällaiset epäselvät tapaukset huomiotta. Olen lisännyt kategoriaan sanat *väki* ja *joukko*, jotka eivät ole minkään sanan monikko-, tai taivutusmuotoja.

RUUMIS ~ KEHO ~ VARTALO

Tämä primitiivi ei nähdäkseni vaadi paljoa käsittelyä. Sen englanninkielinen vastine on BODY. Olen laskenut aineistooni alloleksiksi myös sanan *vartalo*.

(ITSE)

Mainitsen tässä, että Tuovila (2005) on laskenut sanan *itse* kuuluvaksi samaan kategoriaan sanan *minä* kanssa (Virittäjä 2017/2). Olen siitä samaa mieltä, että ITSE tosiaan vaikuttaa huomioimisen arvoiselta, mutta se ei mielestäni kuulu saman semanttisen primitiivin luokkaan kuin MINÄ. Vaikka ajattelen, että sanaa *itse* voisi harkita alkusanojen joukkoon, en ole kuitenkaan ottanut sanaa mukaan varsinaiseen aineistooni.

5.2.2 Muut nominit

JOKIN ~ ASIA ~ (JOKU)

Tästä semanttisesta primitiivistä ei ole paljoakaan mainittavaa. Huomautan lähinnä, että tämä primitiivi viittaa elottomaan asiaan, ja tällöin sana *joku* kuuluu tähän kategoriaan vain puhekielisissä tapauksissa, joissa se ei viittaa ihmiseen: *Siellä oli joku auto*. Primitiivin englanninkielinen vastine on SOMETHING ~ THING.

OSA ~ PALA

Olen lisännyt tähän kategoriaan alloleksiksi sanan *pala*. Valinnastani voidaan olla toistakin mieltä, mutta perustelen tätä seuraavasti: *Osa* viittaa jostakin johonkin asiaan, joka kuuluu johonkin suurempaan kokonaisuuteen, samoin kuin *pala*. *Pala* on lähinnä konkreettisempi siinä missä *osa* voi viitata johonkin abstraktimpaan. Mielestäni kokonaisuuden ja tämän kyseisen primitiivin välinen suhde on kuitenkin sama, minkä vuoksi olen laskenut myös sanan PALA esiintymät aineistooni mukaan.

Sanan englanninkielinen vastine on PART. On hyvä huomata, että edes englanninkieliseen käännökseen ei ole luettu mukaan sanaa *piece*.

TÄMÄ & SE & TUO

Tämä primitiivi on suomen kielen kannalta mielenkiintoinen. Vanhatalo, Tissari ja Idström (2014) ovat päätyneet käännökseen, jossa primitiivin englanninkielistä vastinetta THIS vastaavat suomen kielessä sanat TÄMÄ ja SE. Tämä on mielestäni valintana vielä pohtimista vaativa, sillä sanoilla on

suomen kielessä tärkeä merkitysero. Toisaalta NSM:n piirissä tehdyn tutkimuksen perusteella voitaneen hyvinkin arvioida, että tämä ero eri ole relevantti, mutta kohtelen siltä näitä kahta mainittua sanaa omassa aineistossani erillisinä semanttisina primitiiveinä. Tämän lisäksi lisään vielä kategoriaan sanan *tuo*.

SAMA & ERI

Tämä on ensimmäinen tapaus, joka sisältää yleisemmän huomion. Jos ajatellaan kahden sana olevan semanttisesti toistensa vastakohtia, millä perusteella toinen valitaan semanttiseksi primitiiviksi ja toista ei? Toisen sananhan voi tällöin esittää negaation kautta, tyylisiin *sama* = *ei+eri*. Olen varmuuden vuoksi siis päättänyt laskemaan mukaan molemmat.

Primitiivin englanninkielinen vastine on THE SAME.

SANA

Tämä on toinen primitiivi, jonka alkuperäinen englanninkielinen versio on monikossa, eli sana on englanniksi WORDS; SANAT. Mielestäni tässä pätee sama kuin alkusanan IHMISET kanssa: monikkomuotoisuus ei kuulu nähdäkseni semanttisten primitiivien paradigmaan, tai kääntäen, katson, että alkusanojen tulisi olla lähtökohtaisesti yksikkömuotoisia. Tästä syystä otan sanan aineistooni yksikkömuotoisena, joskin, kuten olen aiemmin kirjoittanut, hyväksyn mukaan myös sen monikkomuotoiset esiintymät.

TOINEN ~ MUU

Tässä semanttisessa primitiivissä on huomionarvoista sanan *toinen* polyseeminen olemus. Tähän kategoriaan en ole laskenut sanaa *toinen* silloin kun se viittaa järjestyslukuun, Tämän primitiivin englanninkielinen vastine on OTHER ~ ELSE.

TOSI ~ TOTTA ~ TODELLA(KIN)

Katson, että viimeinenkin sana voidaan asettaa samaan semanttiseen kategoriaan alkusanan TOSI kanssa. Ero sanan *tosi* ja *todella* välillä on katsoakseni lähinnä syntaktinen. Sanan englanninkielinen vastine on TRUE.

5.2.3 Määrä

YKSI ~ ERÄS & KAKSI

Nämä alkusanat ovat ainoat numeraalit semanttisten primitiivien joukossa. Tätä ovat perustelleet esimerkiksi Tissari ja Vanhatalo (2017) sillä, että periaatteessa mikä tahansa muu luku on esitettävissä näiden kahden luvun avulla.

On syytä mainita, että otan huomioon vain näiden alkusanojen tekstiesiintymät. Tämä johtuu yksinkertaisesti siitä, että numeerinen esiintymä ei sisällä suomen kieltä. Lisäksi suomen kieli eroaa jonkin verran englannista sanan *yksi* käytön suhteen: Englanninkieliset rakenteet, kuten ”I’ll take this one” kääntyvät suomeksi sellaisiksi, että ne eivät sisällä sanaa *one-yksi*: ”Minä otan tämän”.

MUUTAMA ~ PARI

Tämä primitiivi on kenties kyseenalainen. Se ei kuulu sellaisenaan Vanhatalon, Tissarin ja Idströmin (2014) esitykseen, eikä myöskään vuoden 2017 *Virittäjässä* julkaistuun Tissarin ja Vanhatalon listaan alkusanoista, vaan molemmissa katsotaan sanan *muutama* kuuluvan englanninkielisen primitiivin SOME alaiseen kategoriaan. Syy siihen, miksi olen ottanut sen erilleen omaksi kategoriakseen, on erityisesti sana *pari*. Tällä tarkoitan summittaista määrän ilmoittamista, en yksikköä, johon kuuluu kaksi.

Sanaa *pari* voidaan käyttää viittaamaan johonkin tarkkaan määrään (esim. kaksi) tai hieman ympäröivästä ilmaisemaan, että kohteita on jonkin verran, muttei paljon. Sikäli se on tavallaan verrannollinen sanan *muutama* kanssa, ja siksi olen päättänyt tilastoida nämä kaksi samaan kategoriaan erilleen sanoista JOKU ~ JOKIN ~ JONKIN VERRAN

JOKU ~ JOKIN ~ JONKIN VERRAN

Tämä primitiivi vastaa englanninkielistä primitiiviä SOME. Se on sikäli mielenkiintoinen, että leksikaalisen esiintymänsä lisäksi sama merkitys voidaan suomen kielessä ilmaista myös partitiivilla: *He ate some cake – Hän söi kakkua*. Ilmaisun *jonkin verran* käyttäminen tässä yhteydessä on hyvin pitkälti vapaaehtoista.

Katson kuitenkin myös termin *jonkin verran* kuvaavan samaa kohdetta, kuin englannin kielen sana *some*. Tästä syystä olen tilastoinut myös sen esiintymät tähän kategoriaan.

KAIKKI ~ JOKAINEN ~ JOKA

Tämä on primitiivinä jokseenkin selkeä. Olen lisännyt kategoriaan allolekseiksi myös sanat *jokainen* ja *joka*, jotka mielestäni sopivat tämän semanttisen primitiivin alle. Mainitsen, että tässä tilanteessa en tietenkään laske mukaan relatiivipronominia *joka*.

On hyvä huomioida, että tämän primitiivin englanninkielinen vastine on ALL, mutta ei EVERY. Tässä tapauksessa on mielestäni kyse myös konkreettisen ja ei-laskettavan tai abstraktin välillä, ja olen linjakseni asettanut sen, että tämä eri yksin ei riitä sulkemaan jotain sanaa pois semanttisten primitiivien listalta.

PALJON ~ MONI & VÄHÄN ~ HARVA

Näihin primitiiveihin ei ole paljoakaan kommentoitavaa. Huomautan kuitenkin, että näissä primitiiveissa on nähtävissä esimerkiksi edellisessä esimerkissäni käyttämäni linjaero konkreettisen ja ei-laskettavan tai abstraktin välillä. Se ei esittäydy tässä primitiivissä distinktiivisenä piirteenä.

Primitiivien englanninkieliset vastineet ovat MUCH ~ MANY & LITTLE ~ FEW

5.2.4 Adjektiivit

Itkonen (2008) ja Pajunen (luentosarja 2018) kuvaavat kielitypologisen tutkimuksen olevan jonkin verran skeptinen suhteessa adjektiivien universaaliuteen. Näin ollen kaikki adjektiiveiksi laskettavat sanat voivat olla universaaleina kyseenalaisia. Suomen kielessä ne kuitenkin esiintyvät, joten sen perusteella niitä ei voi aineistosta karsia pois.

HYVÄ & PAHA

Nämä sanat eivät juuri kaipaa selitystä. Ne ovat mielestäni adjektiiveina hyvinkin primitiivisiä.

HUONO

Syy tähän lisäykseen on se, että englannin sana *bad* voidaan suomeksi kääntää joko *paha* tai *huono*. Tällöin riippuu kontekstista, kumpi esiintymä on kyseessä. *Virittäjässä* (2017) julkaistussa artikkelissaan Tissari ja Vanhatalo painottavatkin, että alkuperäinen englanninkielinen alkusana BAD tarvitsee vielä tarkennusta. Tässä tutkielmassa lasken myös sanan HUONO esiintymät, mutta teen sen erilleen sanasta PAHA.

ISO ~ SUURI

Sana ovat keskenään allolekseja, vaikka kuten Vanhatalo, Tissari ja Idström (2014) huomauttavat, *suuri* voi sisältää myös sellaisia abstrakteja merkityksiä, joita *iso* ei voi.

PIENI ~ PIKKU

Olen lisännyt tähän kategoriaan semanttiseksi alloleksiksi taipumattoman attribuutin tai yhdyssanan osan *pikku*. Tämän primitiivin englanninkielinen vastine on SMALL, ei esimerkiksi LITTLE.

-LAINEN/-LÄINEN ~ LAJI

Tämä on yksi harvoista primitiiveistä, jotka suomen kielessä esiintyvät pääteaineiksena, ei itsenäisenä lekseiminä. Sen englanninkielinen vastine on KIND. Kokonaisuudessaan se kuvaa ominaisuuden vertailua (*sellainen, tällainen, tuollainen, millainen*), mutta se voi esiintyä myös sanoissa, joissa se on esimerkiksi adjektiivin osa (*laihanlainen*) tai vaikkapa kansallisuuden ilmentäjä (*suomalainen, ruotsalainen*). Olen laskenut sen kaikki esiintymät mukaan aineistooni, sillä vaikka esimerkiksi *suomalainen* käännetään englanniksi *Finnish*, sen voi ilmaista myös tönköllä ilmaisulla *of Finnish kind*.

Tissari ja Vanhatalo (2017) myös katsovat sanan *laji* sisältyvän tämän semanttisen primitiivin kategoriaan. Tästä syystä olen ottanut sen huomioon tilastoinnissani.

5.2.5 Verbit

Kategoriana esimerkiksi Itkonen (2008) kuvaa verbien olevan nykyisen tiedon valossa universaali. Näin ollen on mielestäni perusteltua eritellä verbit omaksi kategoriakseen myös NSM:n primitiivejä kuvattaessa.

AJATELLA ~ MIETTIÄ ~ LUULLA ~ OLLA MIELTÄ

Olen lisännyt tähän joitakuuta allolekseja, mutta tiedostan, että ne eivät ole semanttisesti täysin yhteneviä. Kuten englanninkin vastaava primitiivi THINK voi tarkoittaa eri asioita, niin voi suomenkielinen AJATELLAkin. Yksi suuri ero on siinä, seuraako ko. verbiä konjunktio *että/that*. Jos seuraa, niin puhuja voi verbiä käyttämällä esittää esimerkiksi näkemyksensä jonkin asian tilasta sen sijaan, että hän puhtaasti kuvittelee jonkin entiteetin mielessään. Tästä erosta tietoisena olen päätenyt silti laskemaan kategoriaan mukaan myös hyvin pitkälti tällaista mielipidettä tai näkemystä kuvaavat

verbit *olla mieltä* ja *luulla*. Olen kuitenkin tietoinen, että tämä valintani voidaan myös kyseenalaistaa. Tässä semanttisessa primitiivissä myös noudatan kuvailemaani *kvasiholistisuutta*, jonka perusteella katson, että osat voivat olla erilliset kokonaisuuden kanssa.

TIETÄÄ

Hypoteettisesti laskisin tähän kategoriaan myös sellaisen kopulalauseen kuin *joku on tietoinen*, mutta se ei noussut aineistosta esiin.

TAHTOA ~ HALUTA

Tätä alkusanaa on käsitelty jo tässä tutkielmassa jokseenkin kattavasti. Vanhatalo, Tissari ja Idström (2014) ovat mielestäni esittäneet sille hyvät perustelut. Ulla Vanhatalo on ilmaissut sähköpostiviestissään, että on alkanut taipua pikemminkin sanan HALUTA kannalle tämän semanttisen primitiivin suomenkielisenä edustuksena. Tämä ei varsinaisesti ole, kuten olen ilmaissut, oman työni kannalta relevanttia sen käytännöllisen luonteen vuoksi, mutta kiintoisaa. Asiasta lisää löytyy *Virittäjän* artikkelista (Vanhatalo & Tissari, 2017).

Mainittakoon, että laskisin myös ilmaisun *jonkun tekee mieli* tämän semanttisen primitiivin esiintymäksi, mutta se ei noussut aineistosta esiin.

TUNTEA ~ TUNTUA

Tässä primitiivissä on tärkeää erottaa *tuntea* sanan homonyyminen tai polyseeminen kahtiajako: Suomen kielessä kyseinen verbi voi myös kuvata jonkin asian syvällistä tietämistä, sitä, että ihminen on subjektille tuttu tai sitä, että subjekti on asiaan perehtynyt. Tämän primitiivin osalta olen laskenut aineistooni vain sellaiset esiintymät, jotka kuvaavat aistillista tuntemista. Sanan englanninkielinen vastine on FEEL.

NÄHDÄ & KUULLA

Katsoisin, että käytön ja frekvenssin perusteella aktiiviset verbit *katsoa* ja *kuunnella* olisivat huomattavasti käytännöllisemmät parafraasien muodostamisen perusteella, mutta tämä ei ole nähdäkseni peruste laskea niitä semanttisten primitiivien joukkoon.

SANOA

Alkuperäisessä NSM-alkusanalistassa vain verbi *sanoa* lasketaan semanttiseksi primitiiviksi, ei aineiston perusteella huomattavasti yleisempää verbiä *puhua*. Sikäli ero on kuitenkin ymmärrettävä, että *puhua* vaikuttaa primitiivisyyden sijaan pikemminkin jokseenkin primitiiviseltä molekyyliltä, jonka voi selittää käyttämällä sanaa *puhua*.

TEHDÄ

Olen selkeyden vuoksi jakanut sanan *tehdä* kahtia sen polyseemisen luonteen vuoksi. Englannissa tehdään selkeä ero sanojen *do* ja *make* välillä, mutta suomen kielessä molemmat voidaan ilmaista samalla verbillä. Sanan englanninkielinen vastine on DO, joten en laske valmistamista kuvaavia tekoja aineistooni mukaan. Myös Tissarin ja Vanhatalon huomiot (2017) tukevat tätä valintaa.

TAPAHTUA

Tästä sanasta ei ole paljoa sanottavaa aineistoni puitteissa. On kuitenkin hyvä tiedostaa, että sen käyttö on mahdollisesti rajoittuneempi suomen kielessä kuin englannin kielessä: Suomessa sana ei sisällä polyseemisiä elementtejä, mutta englannissa sillä on: *What happened here? – I happen to be a member of the municipal council.*

Sikälihän *tapahtua* on mielenkiintoinen sana, että se ei normaalisti suomen kielessä esiinny kuin yksikön ja monikon kolmannessa persoonassa.

LIKKUA

Liikkua on verbinä hyvinkin primitiivinen ja sen tilalla esiintyy usein erilaisia deskriptiivisempiä muotoja, kuten *kulkea*, *tulla* ja *mennä*. Nämä muodot eivät kuitenkaan enää ole tämän semanttisen primitiivin esiintymiä. Tämän alkusanan englanninkielinen vastine on MOVE.

On hyvä huomata, että sanan *move* voi suomeksi kääntää ainakin kahdella eri tavalla: *liikkua* ja *liikuttaa*. Vaikka jälkimmäistä en aineistossani ole kohdannut, se on mielestäni ilmaisuna sellainen, joka sisältää myös alkuperäisen *liikkua* verbin. Jos sana ilmaistaisiin englanniksi kiertoilmauksella, sen voisi esittää esimerkiksi muodossa *to make something move*, jolloin tämä ilmaisu sisältäisi myös alkusanan MOVE. Tästä syystä laskisin verbin *liikuttaa* myös mukaan aineistooni.

KOSKEA

Sana *koskea* on polyseeminen. Sen englanninkielinen vastine on TOUCH. Näin ollen en ota aineistooni mukaan sellaisia esiintymiä, jotka kuvaavat joko kivun tunnetta tai sitä, että sana *koskea* viittaa jonkun asian kuulumista jollekulle, kuten esimerkklauseessa *tämä ei koske sinua*. Olen käsitellyt sanaa myös luvussa 2.3.

OLLA

Kenties eniten selvittämistä vaativa osuus menetelmissäni on se, miten olen kohdellun verbiä *olla* – englanninkieliseltä vastineeltaan BE ~ THERE IS ~ HAVE (Vanhatalo, Tissari ja Idström [2014] kirjoittavat että, myöhemmässä NSM-alkusanalistassa HAVE on myös korvattu sanalla BE).

Alkuperäisissä Vanhatalon, Tissarin ja Idströmin (ibid.) esittämissä suomenkielisessä ja englanninkielisessä listassa on molemmissa hajautettu kyseinen semanttinen primitiivi neljään osaan. Nämä osat ovat BE (somewhere) – OLLA (jossain), THERE IS – OLLA (olemassa), HAVE – OLLA (omistaa) ja BE (someone/something) – OLLA (joku/jokin), joita kutsun tästedes samassa järjestyksessä nimillä OLLA1, OLLA2, OLLA3 ja OLLA4.

Esimerkiksi Pajunen (2018) kuitenkin esittää, että suomenkielisellä kopulalauseella on ainakin kuusi funktiota: luokkapredikaatio (mansikka on marja [sic]), ekvaatiopredikaatio (Iltatähti on Aamutähti), ominaisuuspredikaatio (poika on iloinen), lokaatiopredikaatio (poika on pihalla), eksistenssipredikaatio (pihassa on keinu) ja omistuspredikaatio (pojalla on kirja / kirja on pojan). Erittely on muuten sama, mutta ominaisuuspredikaatio, ekvaatiopredikaatio ja luokkapredikaatio on Vanhatalon, Tissarin ja Idströmin (ibid.) artikkelissa niputettu OLLA 4 esiintymän alaisuuteen. Tämä on kenties perusteltua, sillä kopulaverbi *olla* tosiaankin esiintyy samanmuotoisena sekä samanlaisessa syntaktisessä funktiossa kaikissa näistä tapauksista. Lisäksi näin pedantti tarkastelu tuskin on ainakaan omassa työssäni edes relevantti. Jätän siis Pajusen kuusijakoisen kopulan huomion tasolle.

Sen sijaan olen joutunut aineistostani nousseen paineen vuoksi tekemään muita muutoksia. Ensimmäinen alkuperäisestä poikkeava kategoria on tempuksissa perfekti ja pluskvamperfekti esiintyvä *olla*-verbi. Voidaan kenties perustellustikin väittää, että NUT-partisiipin kanssa esiintyvä verbi olisi predikaatiivia indikoiva, mutta katson nämä tempustapaukset omaksi kategoriakseen OLLA5 varmuuden vuoksi.

Seuraava kategoria, OLLA6, on nesessiivirakenteen yhteydessä olevan genetiivisubjektin kanssa esiintyvä *olla*-verbi, kuten esimerkeissä *minun on juostava* ja *Pasin on pakko mennä nukkumaan*. Tämän niputtaminen johonkin edellisistä voisi mielestäni osoittautua sen verran hankalaksi, että otin sen varmuuden vuoksi omaksi kategoriakseen.

OLLA7 kategoria käsittää sellaiset *olla*-verbin infiniittiset tapaukset, joissa niitä ei voida sisällyttää mihinkään edellisistä. Tällaisia ovat esimerkiksi aavistuksen epäkieliopillinen tulevaisuutta predikoiva *hän tulee olemaan suosittu* sekä suurin osa VA-partisiipin ilmentymistä.

Viimeinen numeerinen kategoria, OLLA8, kuvaa sellaisia esiintymiä, joissa *olla*-verbin kategorian määrittäminen on mahdotonta. Paikoin *olla*-verbi esiintyy aineistossa kokonaan syntaktisen kontekstin ulkopuolella yksin, jolloin sen funktiota ei ole aina mielekäästä edes yrittää määrittää.

Kuten mainitsin luvussa 5.1, minulla ei ole myöskään mitään syytä karsia pois *olla*-verbin potentiaaumuotoa *lienee*.

ELÄÄ & KUOLLA

Näiden sanojen kohdalla ainoa huomioni on pohdinta siitä, ovatko nämä tosiaan mahdottomia hajottaa reduktiivista parafrasaa käyttäen. En epäile, etteivätkö ne täyttäisi universaalien sanojen vaatimusta, mutta semanttisen primitiivin toinen kriteeri, sen hajottamattomuus, tuntuu näiden sanojen kohdalla kyseenalaiselta.

VOIDA ~ SAADA

Tämä primitiivi kuvaa mahdollisuutta jonkin asian tekemiseen. Olen lisännyt sen alloleksiksi polyseemisen sanan SAADA, jolla siis tarkoitan tilannetta, jossa vallitsevan tilanteen tai jonkun valta-aseman vuoksi saadaan lupa asian tekemiseen. Perustelen valintaani sillä, että nämä kaksi voivat esiintyä kielessä limittäin. *Voinko mennä uimaan ja saanko mennä uimaan* tarkoittavat jotakuinkin samaa asiaa: Molemmat ottavat huomioon jonkun toisen valta-aseman tähän asiaan liittyen. Toisaalta voidaan ajatella, että sana *voida* on tällä tavalla polyseeminen. Näitä sanoja voisi havainnollistaa esimerkiksi Venn-diagrammilla, jossa kaksi ympyrää leikkaa toisensa.

Toinen ympyröistä kuvaa verbiä *voida* siinä merkityksessä, että tekijällä on kaikki mahdollisuudet jonkun asian suorittamiseen. Toinen ympyrä kuvaa verbiä *saada* siinä merkityksessä, että tekijä saa jotakin jonkun toisen antamana. Leikkaava alue puolestaan kuvaa sellaista tilannetta, jossa nämä kaksi tarkoittavat jotakuinkin samaa, kuten antamassani esimerkissä.



EI ~ ÄLÄ ~ ABESS.

Sana EI on mielestäni primitiivinä hyvinkin perusteltu. Olen myös ottanut mukaan sen imperatiivimuodon ÄLÄ taivutusmuotoineen. Primitiivin englanninkielinen vastine on NOT. Tissari ja Vanhatalo (2017) ovat ottaneet tämän semanttisen primitiivin esiintymäksi myös abessiivimuodot, mistä olen samaa mieltä.

Kiintoisa kysymys mielestäni on kuitenkin se, miksi EI on primitiivi, mutta KYLLÄ ei. Tässä on toisaalta kyse negaatiosta, mutta sikäli se saattaa perustua jonkin laajempaankin: Ensimmäkin kieli on yleensä lähtökohtaisesti positiivista. Jotta voi tehdä negatiivisen lauseen, täytyy erikseen käyttää negaationsanaa tai -rakennetta, eli toisin sanoen kumota vallitseva asiantila. Näin ollen kielen positiivisuus on siis oletettu asiantila, eikä sitä tarvitse erikseen indikoida. Kielitypologiassa pidetään negaatiota hyvinkin universaalina piirteenä (esim. Pajunen 2018)

(KÄYTTÄÄ)

Olen lisännyt tämän verbin mukaan huomiona, sillä se esittäytyy minulle jokseenkin primitiivimäisenä sen suhteen, että sitä vaikuttaa olevan hankala hajottaa pienemmiksi osiksi reduktiivista parafrasaa käyttäen. En ole laskenut sanaa mukaan aineistooni.

(TÄYTYÄ ~ OLLA PAKKO ~ JOUTUA ~ PITÄÄ)

Tähän pätevät samat perusteet kuin edelliseen: Minusta näiden sanojen hajottaminen vaikuttaa hankalalta. Tissari ja Vanhatalo (2017) kuitenkin huomauttavat, että nämä sanat eivät ole universaaleja. En ole laskenut sanaa mukaan aineistooni.

5.2.6 Aika

Tämän kategorian sanojen ongelmat tiivistyvät mielestäni heti sen ensimmäisessä primitiivissä. Kategoria sisältää mielestäni sellaisia sanoja ja rakenteita, jotka eivät ole suomenkielisessä ympäristössä mitenkään primitiivisiä; ne voidaan siis hajottaa pienemmiksi reduktiivista parafrasaa käyttäen.

MILLOIN ~ JOLLOIN ~ SILLOIN/SITTEN ~ AIKA

Ensiksi triviaali huomio: Tässä primitiivissä AIKA ei luonnollisestikaan viittaa sanan *melko* kanssa jotakuinkin synonyymiseen sanaan *aika*.

Tämä primitiivi on kategoriana laaja. Se sisältää paljon erilaisia ajanilmauksia, kuten esimerkiksi kaikki ilmaukset, joissa käytetään sanaa *aika*. Sikäli kategorian laajuus on perusteltua, että esimerkiksi *milloin* sana on ilmaistavissa myös sanoilla *mihin* ja *aikaan*, joskin tällöin se sisältäisi myös sanan *mikä*. Samalla tavalla on hajotettavissa myös kategorian muut sanat: SILLOIN voidaan ilmaista myös *siihen aikaan*, kuten myös lisäämäni alloleksi SITTEN, jolla tarkoitan synonyyminomaisia esiintymiä sanan *silloin* kanssa, en käyttöä, jossa kuvataan kronologista järjestystä: *Mies meni kauppaan ja sitten hän meni baariin*. Myös JOLLOIN tuntuu vahvasti sanalta, joka oikeastaan voidaan hajottaa käyttämällä sanaa *aika* tai muita sanoja: *Ostin oluen, jolloin vaimoni suuttui* → *Ostin oluen, ja silloin (→ ja siihen aikaan/ja sillä hetkellä) vaimoni suuttui*.

Paikoin olisi myös perusteltua sisällyttää tähän kategoriaan sanan *kun*, kuten Tissari ja Vanhatalo (2017) esittävät. Olen kuitenkin päätenyt laskemaan sen omaksi esiintymäkseen, mistä lisää myöhemmin.

Kyseenalaistan siis hieman tämän kategorian laajuuden, mutta ennen kuin asiaan perehdytään paremmin, en tee sen perusteella johtopäätöksiä. Olen laskenut Vanhatalo, Tissari ja Idström (2014) ehdottamat sanat mukaan kategoriaan. Primitiivin englanninkielinen vastine on WHEN ~ TIME.

NYT ~ HETI ~ TÄLLÄ HETKELLÄ

Tähän olen lisännyt kaksi ehdotusta allolekseiksi. Sana englanninkielinen vastine on NOW.

ENNEN ~ AIKAA SITTEN ~ AIKAISEMMIN & ENNEN (JOTAKIN)

Tämä ei ole mielestäni mitenkään ongelmallinen. Olen kuitenkin eritellyt omiksi kategorioikseen sanat, jotka viittaavat menneeseen aikaan yleisesti ja sanat, jotka selkeästi ilmaisevat jonkin kiintopisteen, jota edeltävästä ajasta puhutaan. Primitiivin englanninkielinen vastine on BEFORE.

JÄLKEEN ~ AJAN KULUTTUA ~ MYÖHEMMIN ~ SITTEN ~ PÄÄSTÄ

Tähän kategoriaan olen lisännyt pari ehdotusta allolekseiksi.

KAUAN (AIKAA) ~ PITKÄ AIKA & VÄHÄN AIKAA & JONKIN AIKAA

Nämä kategoriat ovat mielestäni erikoisia. Miksi ne lasketaan omiksi primitiiveikseen, kun ne selkeästi voidaan hajottaa reduktiivista parafrasaa käyttämällä pienempiin osiin? Myönnän kyllä, että

KAUAN on leksikaalistunut muotoon, joka ei sisällä sanaa *aika*, mutta nähdäkseni se voidaan silti ilmaista parafrasilla.

Vaikka kyseenalaistan kaikki nämä primitiivit, olen silti laskenut ne mukaan aineistooni. Lähinnä haluan painottaa niiden ongelmallisuutta. Sanojen englanninkieliset vastineet ovat A LONG TIME, A SHORT TIME ja FOR SOME TIME.

HETKI

Tämä on mielestäni perusteltu primitiivi. Sen englanninkielinen vastine on MOMENT.

(JOSKUS ~ VÄLILLÄ ~ VÄHÄN VÄLIÄ & USEIN & KOKO AJAN & AINA)

Olen tarkastellut aineistostani yhteydessä erilaisia ajanmääreitä. Ne eivät mielestäni täytä välttämättä primitiivin kriteerejä, vaan ovat pikemminkin hyödyllisiä molekyyilejä. En kuitenkaan lue niitä mukaan aineistooni.

(VIELÄ & VIELÄ & YHÄ)

Olen huomioinut aineistostani polyseemisen sanan *vielä* esiintymiä. Katson, että se on vähintäänkin hankala hajottaa reduktiivista parafrasiasia käyttämällä. Sama pätee sanaan *yhä*.

Sanan *vielä* polyseemisyys osoittautuu ilmauksissa *En ole vielä saanut palkkaa* ja *Minä olen vielä baarissa*. Ensimmäisessä sana kuvaa tapahtumaa, joka ei ole tapahtunut, mutta jonka oletetaan tapahtuvan. Jälkimmäisessä esimerkissä kuvataan asiaa, joka jatkuu, vaikka sen kenties oletettiin jo loppuneen, jolloin sana on siis synonyyminen sanan *yhä* kanssa. En ole laskenut sanoja aineistooni.

(KERTA)

Tämä on yksi ehdotukseni suomen kielessä alkusanamaisena esiintyvälle primitiiville. Sikäli sana voi olla ongelmallinen, että sen englanninkielinen vastine TIME on vahvasti polyseeminen, kuten edellisestä kohdasta käykin ilmi. En ole kuitenkaan laskenut sanaa aineistooni.

5.2.7 Paikka

NSM:n paikankuvaukset sisältävät mm. sellaisia sanoja, jotka kuvaavat kielitypologiassakin universaaleina pidettyjä ilmaisuja (ks. esim. Itkonen 2008, Pajunen 2018). Sikäli katson, että kategoriana paikka on erittäin relevantti universaaliutta käsiteltäessä. Kategorია ei myöskään sisällä suurempia ongelmia.

MISSÄ ~ JOSSA/JOLLA ~ PAIKKA ~ INESS/ADESS

Samaan tapaan kuin ajanilmauksia kuvaavien sanojen kategorian ensimmäinen primitiivi MILLOIN ~ JOLLOIN ~ SILLOIN/SITTEN ~ AIKA, tätäkin voidaan pitää laajana. Toisaalta, tämä primitiivi, jonka englanninkielinen vastine on WHERE ~ PLACE, on mielestäni kuitenkin selkeärajaisempi. Paikanilmauksissa ei myöskään ole sitä ongelmaa, että suuri osa niistä voitaisiin hajottaa reduktiivista parafrasaa käyttämällä.

Olen lisännyt tähän kategoriaan vain kieliopillisen alloleksin JOLLA, sillä suomen kielen vaihtelu sisä- ja ulkopaikallissijojen välillä saattaa olla lähes mielivaltaista vailla mitään eroa merkityksessä (esimerkiksi *keikalla/konsertissa*).

Lisäksi *Virittäjässä* (2017/2) esiintyneessä artikkelissaan Tissari ja Vanhatalo nimeävän myös inessiivin tämän semanttisen primitiivin esiintymäksi. Tämä on mielestäni perusteltua. Asia kaipaakin kuitenkin lähempää tarkastelua. Ensinnäkin inessiivin ja adessiivin vaihtelu on suomen kielessä paikoin jopa mielivaltaista: Erityisesti paikannimissä on sellaista vaihtelua sisä- ja ulkopaikallissijojen välillä, ettei sitä voi ainakaan täysin kieli-intuition mukaan selittää. Miksi joku asuu Pirkkalassa, mutta toinen asuu Tampereella?

Tämän lisäksi ei mielestäni ole aivan selkeää, että tätä semanttista primitiiviä kuvaavat pääätteet jäisivät vain inessiiviin ja adessiiviin. Esimerkiksi elatiivipääte sisältää ajatuksen inessiivistä: Lause *me tulemme konsertista* johtaa oletukseen, että puhuja on ollut *konsertissa*. Onko raja siis esimerkiksi inessiivin ja elatiivin välillä semanttisesti niin suuri, että vain inessiiviesiintymät voidaan laskea tämän semanttisen primitiivin esiintymiksi? Lisäksi, kuten olen kirjoittanut, suhtautumiseni sijataivutukseen on salliva, mikä herättää kysymyksiä siitä, miten vain yhden sijamuodon voisi hyväksyä tietyn semanttisen primitiivin esiintymäksi. Esimerkiksi tilastoidessani sanan TÄSSÄ esiintymiä olen laskenut mukaan myös elatiivi- ja illatiiviesiintymät.

Mielestäni on tärkeää ottaa huomioon myös paikallissijojen kieliopillisuus. Siinä, missä inessiivi on käyttönsä puolesta jokseenkin leksikaalinen, esimerkiksi adessiivilla voidaan jo ilmaista sellaisia asioita, joiden leksikaalisuus on mielestäni vähäisempää kuin puhtaan lokaation ilmaiseminen: Adessiivilla voidaan esimerkiksi ilmaista jonkin käyttämistä välineenä, ja jopa saada

näin aikaan marginaalisia homonyymisiä kaksoismerkityksiä: *Mies surmasi toisen nakkikioskilla*. Lisäksi suomen kielessä on tavallista ilmaista omistusta adessiivipäätteen avulla: *Minulla on olutta*. Sisältääkö tämä ilmaisu lokaation ilmauksen vai ei?

Kieliopillisuus pätee vahvasti myös sisäpaikallissijoihin. Jotkin verbit esimerkiksi vaativat kohteensa tiettyyn sijamuotoon saadakseen tietyn merkityksen: *Minä uskon Jumalaan* tai *minä pidän oluesta*. Tällöin ilmaisussa esiintyvän sijapäätteen kieliopillisuus tulee ottaa huolellisen tarkastelun alaiseksi.

Näistä syistä en poikkeuksellisesti ole laskenut aineistooni mukaan sanojen inessiivi-esiintymiä, vaikka hyväksynkin niiden edustavan semanttista primitiiviä. Asia kaipaa mielestäni vielä tarkempaa pohdintaa. Intuitiivisen kieliälyni sekä aineistoni perusteella voin kuitenkin yleisen maininnan tasolla todeta, että inessiivi on hyvin yleinen muoto suomen kielessä.

TÄSSÄ ~ TÄÄLLÄ

Tämä on hyvin pitkälti ongelmaton primitiivi. Nähdäkseni ainoa huomionarvoinen asia on se, että sanan englanninkielinen vastine HERE voidaan kääntää suomeksi joko *tässä* tai *täällä*, jolloin ero tehdään lähinnä kyseessä olevan paikan perusteella. Eroja sanojen välillä selitetään esimerkiksi kohteena olevan paikan koolla, jolloin *tässä* viittaisi pienempään ja *täällä* viittaisi suurempaan kohteeseen (ks. esim. <https://uusikielemme.livejournal.com/5206.html> 1.4.2018)

On kuitenkin hyvä huomata, että sana *tässä* on homonyyminen. Se voi kuvata edellisessä kappaleessa kuvaamallani tavalla paikkaa, mutta se voi olla myös sanan *tämä* inessiivimuoto. Luvun alussa kuvaamani holistisuusperiaatteen mukaisesti katson, että tällöin sana *tässä* on itseasiassa kahden semanttisen primitiivin esiintymä: Se sisältää primitiivin TÄMÄ, mutta myös inessiivipäätteen, joka tekee siitä samanaikaisesti sanan MISSÄ ~ JOSSA/JOLLA ~ PAIKKA ~ INESS. esiintymän. Sen sijaan se ei sisällä primitiiviä TÄSSÄ ~ TÄÄLLÄ.

SIELLÄ ~ SIINÄ & TUOLLA ~ TUOSSA

Suomen kielessä on nähdäkseni relevanttia erottaa toisistaan sanat *tässä*, *tuossa* ja *siinä* sekä *täällä*, *tuolla* ja *siellä*. En väitä, että ero olisi universaali, mutta suomen kielessä se viestii niin oleellisia asioita puhujan ja kohteen välisestä suhteesta, että olen aineistossani laskenut ne omiksi primitiiveikseen. Sen sijaan eron sanojen SIELLÄ ja SIINÄ sekä TUOLLA ja TUOSSA katson olevan sama kuin edellisessä esimerkissä.

PÄÄLLÄ ~ YLLÄ ~ YLÄPUOLELLA & ALLA ~ ALAPUOLELLA

Nämä sanat kuvaavat sellaisia käsitteitä, joiden kuvataan kielitypologiassa olevan jokseenkin universaaleja (ks. esim. Itkonen 2008, Pajunen 2018). Olen lisännyt molempiin tilastoitaviksi allokseiksi myös sanojen hieman abstraktimmat esiintymät *yllä*, *yläpuolella* ja *alapuolella*. Seija Tuovilan (2005) valinnat tukevat tätä ratkaisua. Sanojen englanninkieliset vastineet ovat ABOVE ja BELOW.

KAUKANA & LÄHELLÄ

Mielestäni molemmat ovat perusteltuja, joskin voisin kuvitella, että ne pystyttäisiin hajottamaan reduktiivista parafrasaa käyttäen myös pienempiin osiin käyttäen esimerkiksi molekyylä *matka*. Sanojen englanninkieliset vastineet ovat FAR ja NEAR.

PUOLI

Tämä sana on suomen kielessä polyseeminen, sillä se voi viitata joko monijakoiseen asiaan, jonka jokin puoli on tarkastelun kohteena, tai se voi viitata siihen, että joku objekti on jaettu tasan kahtia, ja tarkastellaan toista yhtä suurista puolikkaista. Sanan englanninkielinen vastine on SIDE, eli en ole laskenut jälkimmäiseksi esittämiäni esiintymiä aineistooni.

SISÄ- & ULKO-

Primitiivi on SISÄ-, jonka englanninkielinen vastine on INSIDE, on mielestäni perusteltu, mutta se herättää kysymyksen siitä, miksi negaation kautta aikaansaatu sana *ulko-* ei ole primitiivin asemassa. Olen erottanut sanat omiksi kategorioikseen aineistossani.

5.2.8 Kieliopilliset sanat

KOSKA ~ VUOKSI ~ TAKIA ~ SIKSI ~ SILLÄ ~ SIITÄ SYYSTÄ

Kokonaisuudessaan tämä primitiivi kuvaa kausaliteettisanoja. Olen tästä syystä lisännyt sille allokseiksi SIKSI (ei sanan *se* yksikön translatiivi), SILLÄ (ei sanan *se* adessiivi) ja SIITÄ SYYSTÄ.

JOS & KUN

Näiden sanojen välinen ero on mielenkiintoinen. Suomen kielessä se on nähdäkseni selkeä: *Jos* kuvaa mahdollisuutta tai konditionaalisuutta, *kun* kuvaa varmuutta. Tiedostan kuitenkin, että näin ei kaikissa kielissä ole: Esimerkiksi englannissa sanan JOS vastine IF voi myös indikoida esimerkiksi

kysyvää sivulausetta: *I don't know if I should drink beer*. Myös esimerkiksi tšekin kielessä ero sanojen välillä on erilainen kuin suomessa: Sanan *jos* käännös *jestli* voi esimerkiksi indikoida kysyvää sivulausetta kuten englannissa. Tämän lisäksi tosin suomen kielen sanaa *jos* voi paikoin myös vastata *když*, joka normaalisti vastaa suomen kielen sanaa *kun*.

Olen joka tapauksessa ottanut erikseen aineistoni tilastointiin mukaan sanan KUN. Katson, että se on suomen kielessä relevantti, joskin se voitaisiin, kuten sanottua, liittää primitiiviin MILLOIN ~ JOLLOIN ~ SILLOIN/SITTEN ~ AIKA.

5.2.9 Taipumattomat sanat

Tähän kategoriaan olen aineistoni selkiyttämiseksi sijoittanut sellaiset taipumattomat tai vaillinaisesti taipuvat sanat, jotka eivät sovi muihin kategorioihin

ENEMMÄN ~ LISÄÄ ~ ENÄÄ

Tämä primitiivi on mielenkiintoinen, sillä se on komparatiivimuoto toisesta primitiivistä, primitiivistä PALJON ~ MONI. Miksi tämä siis on primitiivi, mutta esimerkiksi *vähemmän* ei? Kyseessä on mahdollisesti vastakohtaisuus tai negaatio, mutta mielestäni tilanne on epätasapainoinen.

EHKÄ ~ KENTIES

Tähän primitiiviin olen lisännyt alloleksiksi sanan KENTIES. Voidaan tosin sanoa, että sanana *ehkä* on neutraalimpi kuin *kenties*. Tässä työssä kuitenkin kohtelen sitä synonyyminomaisena alloleksinä. Primitiivin englanninkielinen vastine on MAYBE.

NÄIN ~ KUTEN ~ NIIN ~ TAVALLA

Tähän ei lasketa homonymisen sanan *näin* merkitystä yksikön ensimmäisen persoonan aktiivin imperfektinä. Lisäksi on hyvä tehdä ero sanan NIIN kolmen eri merkityksen osalta: Sana voi kuvata puhunnan kanssa yhtä mieltä olemista (*Nyt on kaunis ilma – Niin*), se voi olla jotakuinkin synonyyminen ilmauksen *sillä tavalla* kanssa (*Ei se ole niin!*) tai se voi toimia painotuksena esimerkiksi jonkin asian ominaisuutta kuvattaessa (*Hän on niin kaunis!*). Olen aineistooni laskenut kaksi ensimmäistä vaihtoehtoa, mutta en viimeistä. Sanan englanninkielinen vastine on LIKE ~ AS ~ WAY.

Olen myös ottanut mukaan aineistooni sanan *tapa* adessiiviesiintymät silloin, kun ne kuvaavat jonkin tekemistä *tietyllä tavalla*. Tämä ratkaisu noudattaa Tissarin ja Vanhatalon linjausta (2017).

HYVIN ~ ERITTÄIN ~ TOSI ~ KOVIN ~ TODELLA

Tähän primitiiviin olen lisännyt vain alloleksejä, jotka otan huomioon tilastoinnissani. Primitiivin englanninkielinen vastine on VERY. Mainittakoon, että *hyvin* on homonyyminen sana. Tässä tapauksessa tarkoitan sen käyttöä vahvistus- tai tehostussanana, esim. *Minä rakastan sinua hyvin paljon*.

(AIKA ~ MELKO ~ IHAN)

Esittelen tämän edelliseen primitiiviin liittyen. Mielestäni sen olemus on sellainen, että se voisi olla alkusana, mutten kuitenkaan ole ottanut sitä mukaan aineistooni. Se ei myöskään vaikuta helposti hajotettavalta reduktiivista parafrasaa käyttäen.

(KANSSA ~ LUONA ~ MUKANA & ILMAN)

Myös tämän hajottaminen reduktiivista parafrasaa käyttämällä vaikuttaa hankalalta. Kenties on mahdollista muotoilla tämä ehdottamani primitiivi esimerkiksi lauseella, jossa käytetään rinnastuskonjunktia, mutta saadaanko tällaisella menettelyllä aikaiseksi sama vaikutelma, joka tätä kyseistä sanaa käytettäessä syntyy? Oleellista tässä sanassa ei ole pelkästään se, että kaksi asiaa ovat tilanteessa samaan aikaan, vaan myös se, että ne ovat vuorovaikutuksessa toistensa kanssa. En lue näitä sanoja mukaan aineistooni.

(ENSI ~ SEURAAVA & VIIME ~ EDELLINEN)

Nämä sanat kuvaavat hyvin pitkälti asioiden kronologista järjestystä, joka voi olla hyvinkin relevantti tekijä eksplikaatiota kirjoitettaessa. Kenties tällaisen järjestyksen voisi kuvailla myös käyttämällä alkusanaa SITTEEN, mutta mielestäni se vaikuttaa vähintäänkin ongelmalliselta.

Lisäksi näiden sanojen hajottaminen reduktiivista parafrasaa käyttämällä vaikuttaa mielestäni hankalalta. En kuitenkaan lue näitä sanoja mukaan aineistooni.

(KOKO)

Olen huomionut tämän sanan aineistostani siksi, että se painottaa samaa osa-kokonaisuus-jatkumoa kuin alkusana OSA. Sikäli sanan *koko* voisi ilmaista negatiota ja sanaa *osa* käyttämällä, mutta näin saatettaisiin päätyä jopa kehämääritelmäiseen eksplikaatioon. En kuitenkaan lue tätä sanaa mukaan aineistooni.

5.2.10 Kysymyssanat – epävirallinen kategoria

Kategoriana kysymyssanat ovat mielenkiintoinen. En ole ottanut niitä mukaan varsinaiseen aineistooni, mutta mielestäni ne sisältävän merkitystä, joka on joko vaikea tai mahdoton hajottaa reduktiivista parafraasia käyttäen. Niiden universaaliutta en kuitenkaan voi testata.

(KUKA)

Tämä on ensimmäinen aineistostani esiin noussut kysymyssana. Se on mielestäni hankala hajottaa reduktiivista parafraasia käyttäen.

(MIKÄ)

Tarkoitan tällä nimenomaan kysymyssanaa *mikä*, en relatiivipronominia. Tämä kysymyssana on sikäli mielenkiintoinen, että katson sen olevan eräänlainen kysymyssanojen semanttinen primitiivi: Sitä käyttämällä voidaan muodostaa parafraasi monelle muulle kysymyssanalla: *miten* → *millä tavalla*.

(KUINKA)

Kysymyssanoja koskevat perustelun pätevät myös tähän sanaan. Katson sanan olevan oma erillinen kysymyssanansa.

(MIKSI)

Tämä on kysymyssanoista viimeinen, jonka olen huomioinut erikseen. Voidaan tosin väittää ihan perustellustikin, että tämä sana voidaan muotoilla parafraasilla käyttämällä kysymyssanaa *mikä*: *miksi* → *mistä syystä*. Pidän sanaa silti niin kiteytyneenä, tarkastelen sitä omana sananaan. Kuten luvun alussa olen tosin sanonut, halutessaan kuka tahansa voi suorittaa tilastoinnin päättäen asian toisin.

6. TULOKSET

Tässä luvussa pyrin esittelemään oleelliset tulokset, jotka ovat aineistostani nousseet. Etenen yleisimmistä huomioista yksityiskohtaisempiin. Luku koostuu neljästä alaluvusta.

Aloitan esittämällä, kuinka suuri määrä koko aineiston sanoista koostuu semanttisten primitiivien esiintymistä ts. alkusanoista. Esitän tämän sekä koko kirjasta että osa-alueiden (tekstit, tehtävät, sanastolliset kokonaisuudet) osalta. Näin tekemällä pyrin antamaan yleiskuvan siitä, millainen NSM-alkusanojen kerrostuma on oppikirjakontekstissa. Havainnollistan myös alkusanojen jakaumaa kirjassa: Jotkin sanat ovat erittäin yleisiä, toiset eivät esiintyneet aineistossa kertaakaan. Esitän alkusanojen määrän aineistossani pylväsdiagrammeina, joista näkyy selvästi, mitkä sanat ovat harvinaisia ja mitkä erittäin yleisiä. Kahdeksi ääripääksi olen ottanut sanat, jotka eivät tosiaan esiinny aineistossa kertaakaan sekä ne sanat, joilla on esiintymiä yli sata. Erityisesti näitä ääripäitä vertaan suomen kielen taajuussanastoon.

Seuraavassa alaluvussa jaan aineistoni osa-alueiden mukaan. Näin voin tarkastella sitä, ovatko alkusanat yleisempiä esimerkiksi teksteissä kuin tehtäväosioissa.

Kolmannessa alaluvussa pyrin vertailemaan sitä, miten alkusanojen määrä muuttuu kirjan edetessä: Onko loppupään kappaleissa enemmän alkusanoja kuin alkupään kappaleissa? Teen tämän esittämällä sekä alkusanojen kokonaismäärän että suhteellisen määrän, ja näiden tilalle tuon kolmannen käsitteen: varianssin. Tämä tarkoittaa sitä, että mitä suurempi luvun varianssilukema on, sitä useampi alkusana on siinä edustettuna.

Viimeisessä alaluvussa tarkastelen alkusanoja yksityiskohtaisemmin: Esitän luvussa 5 käyttämieni kategorioiden pohjalta diagrammit, jotka havainnollistavat sitä, kuinka tietty alkusana on edustettuna aineistossa. Lopuksi tarkastelen vielä sitä, miten mikäkin alkusana on edustettu mil-läkin kirjan osa-alueella, ja teen myös jokaisen kappaleen top3-listan.

6.1 Yleiskuva NSM-sanoista kirjoissa

Koko aineistoni sisältää yhteensä 17 276 sanaa. Näistä alkusanoiksi katsomiani esiintymiä löytyy 4119, eli 23,8%. Kuten olen esitellyt, olen jakanut aineistoni kolmeen osa-alueeseen: teksteihin, tehtäviin ja erityisiin sanastoa opettaviin kokonaisuuksiin. Näistä kahdesta tekstiä ja tehtäviä ovat kokonaissanamäärältään jotakuinkin samankokoiset: tekstiosa-alueiden kokonaissanamäärä on 7632

ja tehtävien 7575. Alkusanoja tekstiosa-alueissa on 1896 ja tehtävissä 1677. Kolmas osa-alue, sanastokokonaisuudet, on sanamäärältään selvästi pienempi kuin kaksi muuta; se sisältää 2069 sanaa, joista alkusanoja on 546. Esitän saman datan kaaviolla:

Taulukko 1 – alkusanojen suhde kirjan kokonaissanamäärään

	Kokonaissanamäärä	Alkusanojen määrä	Prosentuaalinen suhde
Koko kirja	17 276	4119	23,8%
Tekstit	7632	1896	24,8
Tehtävät	7575	1677	22,1%
Sanastokokonaisuudet	2069	546	26,4

Taulukosta siis nähdään, että mikään osa-alue ei poikkea kovin suuresti alkusanojen suhteellisen määrän perusteella muista. Sanastokokonaisuudet sisältävät selkeästi vähiten sanoja, mutta alkusanojen suhteellinen määrä niissä on korkeampi kuin muissa osa-alueissa tai koko kirjassa. Marginaali on kuitenkin pieni, joten eri ei ole nähdäkseni järin merkittävä. Tehtävät puolestaan sisältävät prosentuaalisesti vähiten alkusanoja, mutta kuten sanaston osalta, tämäkin ero on koko kirjaan suhteutettuna melko pieni. Ero sanastollisten kokonaisuuksien ja tehtävien välillä kuinkin on 4,3 prosenttiyksikköä, mikä on jo huomionarvoinen tieto.

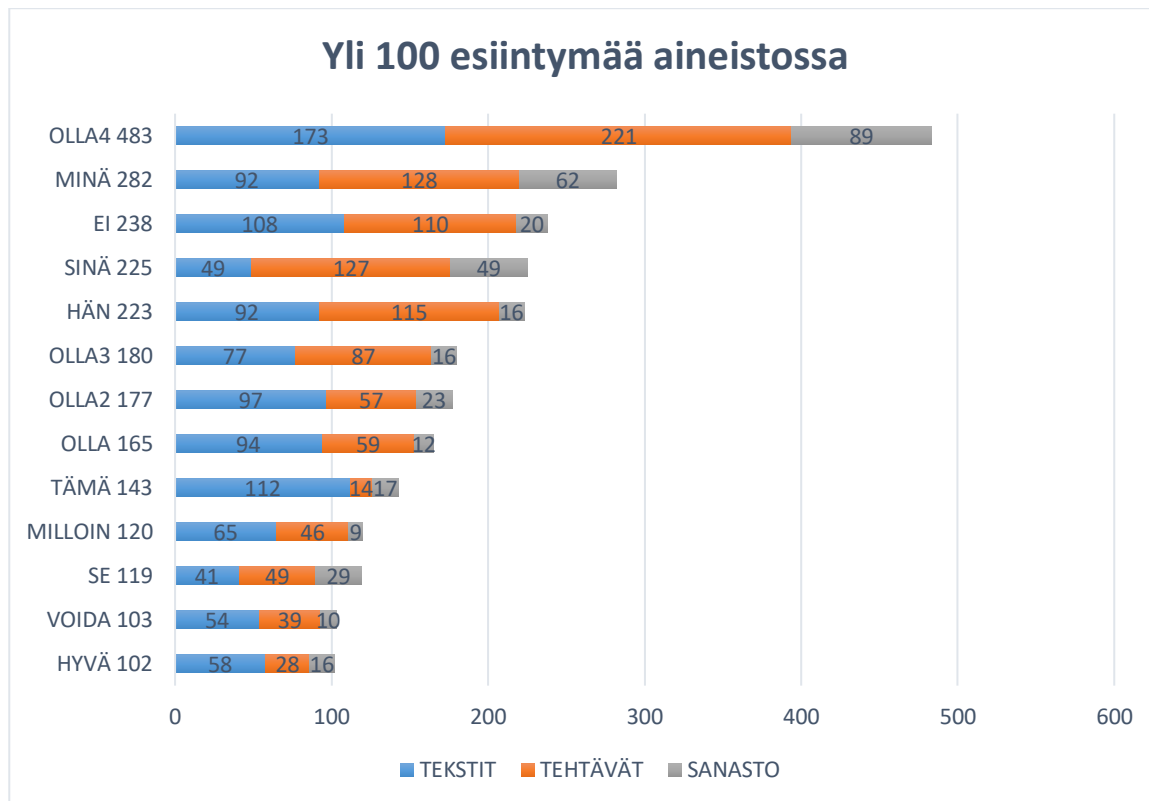
On syytä jälleen painottaa sitä, että kuten menetelmiäni kuvaavassa luvussa 5 selitin, ilmaus saattaa joskus sisältää enemmän alkusanoja kuin sanoja: Sana *tässä* silloin, kun kyseessä on sanan *tämä* inessiivimuoto, sisältää sekä alkusanan TÄMÄ että ehdotetun esiintymän alkusanalle PAIKKA ym., joka siis esiintyy inessiivimuotona. Inessiivimuodoista olen kirjoittanut jo luvussa 5.2. Tämän huomion merkitys on se, että kaikkien sanojen ja alkusanojen suhde ei aina ole selkeä. On kuitenkin syytä myös lisätä, että tällaiset esiintymät ovat aineistossani hyvin harvinaista, enkä katso niillä olevan mainittavaa tilastollista relevanssia.

Seuraavaksi esitän palkkidiagrammein kunkin alkusanan määrän aineistossani. Diagrammit sisältävät vain kyseisten alkusanojen *edussanan*, sillä en ole tilanpuutteen vuoksi voinut kirjoittaa esimerkiksi alkusanaa MISSÄ ~ JOSSA/JOLLA ~ PAIKKA kokonaisuudessaan. Diagrammit siis sisältävät kutakin alkusanaa mielestäni kuvaavimman edussanan, tässä tapauksessa MISSÄ.

Vertailen esittämieni diagrammien lukuja Pauli Saukkosen, Marjatta Haipuksen, Antero Niemikosken ja Helena Sulkalan Suomen kielen taajuussanastoon (1979) ja esittelen merkittävät poikkeukset. Lisäksi kiinnitän huomiota huomattavan erikokoisiin lukuihin sellaisten sanojen

kanssa, jotka kuuluvat esimerkiksi samaan kategoriaan tai ovat muuten semanttisessa suhteessa toisiinsa (esim. *hyvä* ja *paha/huono*).

Diagrammi 1 – alkusanojen määrä aineistossa 1/5



Diagrammista on nähtävissä kaikki alkusanat, jotka esiintyvät aineistossa yli 100 kertaa. OLLA4, joka siis kuvaa jonkin luonteen indikoimista esimerkiksi predikatiivin avulla, on aineiston selkeästi yleisin sana. Myös alkusanat OLLA3, OLLA2 ja OLLA1 ovat aineistossa erittäin yleisiä. Näin voidaan todeta, että *olla*-verbi itsessään on oppikirjan selkeästi yleisimpiä lekseemejä.

Myös aineistossa olleet persoonapronominit, joista siis HÄN on samaa kategoriaa kuin esimerkiksi IHMINEN, ovat erittäin yleisiä. Sama pätee negatiosanoihin, yleiseen aikaa kuvaavaan kategoriaan MILLOIN ~ JOLLOIN ~ SILLOIN/SITTEN ~ AIKA sekä pronomineihin TÄMÄ ja SE. Ainoa Adjektiivi, jolla on aineistossa yli sata esiintymää, on HYVÄ, ja ainoa verbi on VOIDA.

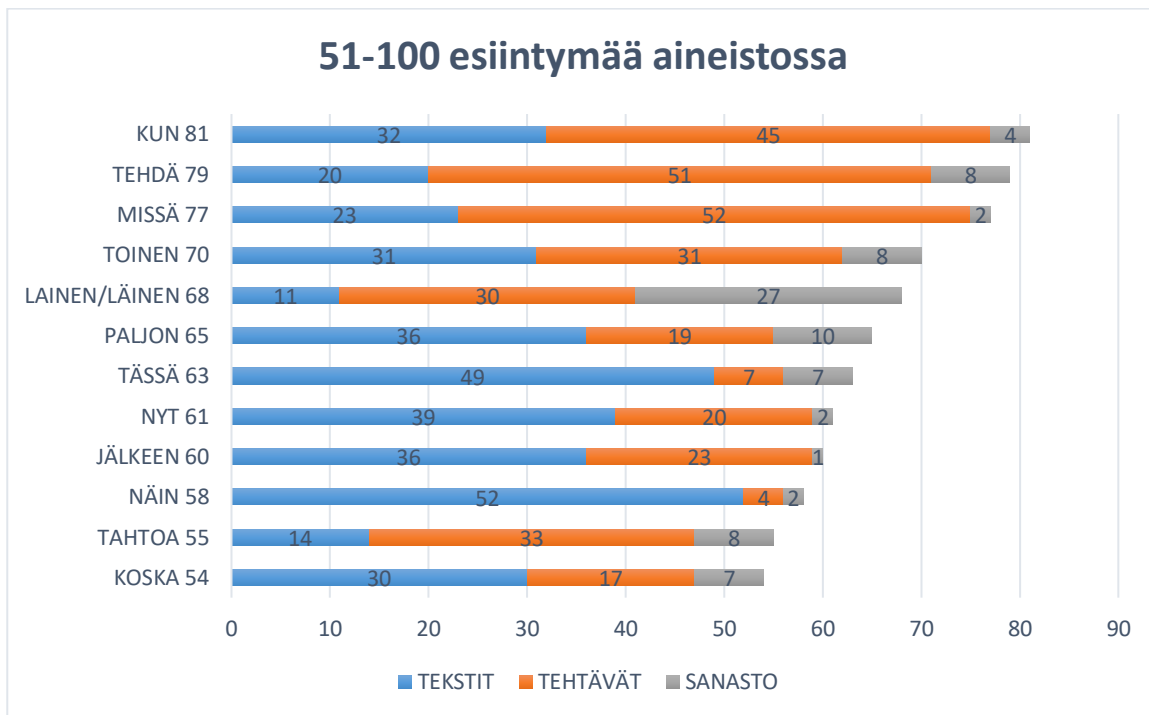
Havainnot noudattelevat kohtalaisen hyvin suomen kielen sanaston kokonaiskuvaa, jota on kuvailtu esimerkiksi Saukkosen ym. (1979) teoksessa Suomen kielen taajuussanasto: Em. esittävät *olla*-verbin todellakin olevan suomen kielen yleisin sana. Myös *se*, *ei*, *tämä*, *hän* ja *voida*

ovat em. mukaan kymmenen yleisimmän suomen kielen sana joukossa. Sen sijaan persoonapronominit *minä* ja *sinä* sekä alkusanan MILLOIN kuvaamat aikailmaisut eivät ole em. mukaan aivan näin yleisiä: *Minä* on vasta suomen kielen 27:nneksi yleisin sana, *sinä* 63:nneksi ja ensimmäinen aikaa kuvaava sana em. Suomen kielen taajuussanastossa on *aika*, joka on sijalla 20. Mielestäni tämä havainnollistaa sitä, että oppikirjat eivät suoranaisesti vastaa luonnollista kieltä. Ne on tehty tiettyä tarkoitusta varten, jolloin myös niiden sisältämä sanasto on pyritty rakentamaan siten, että se tukee tätä päämäärää. Persoonapronominit *minä* ja *sinä* ovat mielestäni siis opetustarkoitukseen itsestäänselvää sanastoa, kuten myös aikailmaukset. Lisäksi persoonapronominien yleisyyttä aineistossani voi selittää sillä, että se esiintyvät jotakuinkin aina myös verbilauseita ja -ilmaisuja opetettaessa.

Sanan HYVÄ yleisyys on mielenkiintoista, joskaan ei yllättävää. Kuten allaolevista diagrammeista käy ilmi, sen vastakkaiset adjektiivit *paha* ja *huono* ovat harvinaisia: *Paha* esiintyy aineistossa vain kolme kertaa ja *huono* vain kuusi. Sanalla *hyvä* puolestaan on esiintymiä 102. Katson, että tämä voi kertoa esimerkiksi oppikirjatekstin pyrkimyksestä esiintyä positiivisena negatiivisen sijaan, mutta tämä vaatii mahdollisesti tarkempaa tarkastelua.

Diagrammi myös tukee silmämääräisesti sitä havaintoa, että sanastolliset kokonaisuudet sisältävät määrällisesti vähemmän alkusanoja kuin kaksi muuta. Jokainen näistä alkusanoista on kuitenkin selkeästi edustettuna myös niissä.

Diagrammi 2 – alkusanojen määrä aineistossa 2/5



Tämä diagrammi kuvaa alkusanoja, jotka esiintyvät aineistossa yli viisikymmentä mutta alle sata kertaa. Näistä eniten esiintymiä on sanalla *kun*, joka, kuten luvussa 5 kirjoitin, voisi olla myös osa kategoriata MILLOIN ~ JOLLOIN ~ SILLOIN/SITTEN ~ AIKA. Se, että sana *kun* on itsessäänkin yleinen kuten mainittu kategoria, tekee tästä ajanilmauksista hyvin yleisiä aineistossani. Myös paikanilmaukset, joiden edussanana toimii siis sana *MISSÄ* ovat aineistossani yleisiä, ja jos tähän kategoriaan sisällytettäisiin inessiiviesiintymät, ne olisivat varmasti vielä paljon yleisempiä. Paikanilmauksista myös *tässä*, joka kuuluu samaan kategoriaan sanan *täällä* kanssa, on aineistossani yleinen.

Lisäksi tähän diagrammiin kuuluu kaksi verbiä, *tehdä* ja *tahtoa*, kaltaisuutta tai johonkin kuuluvuutta kuvaava pääte *-lainen*, ajanilmaus *nyt*, sanan *muu* kanssa samaan kategoriaan kuuluva sana *toinen*, sekä ajan että paikanilmauksena toimivat sanat *jälkeen*. Myös sanat *paljon*, *näin* ja kausaliteettisanoja kuvaava *koska* ovat aineistossani yleisiä.

Suomen kielen taajuussanastoon (1979) verraten kiintoisaa on huomata, että sana *paljon* on aineistossani näinkin korkealla. Sen yleisin edustus *moni* on em. teoksen mukaan vasta suomen kielen 57:nneksi yleisin sana, mistä tosin ei voi vielä vetää järin suuria johtopäätöksiä, sillä aineistoni ulkopuolelle jäi monta yleistä sanaa, jotka eivät täytenä alkusanojen kriteereitä. Samoin perusteiden voidaan kuitenkin päätettävä *-lainen* pitää mielenkiintoisena, sillä ensimmäinen sen sisältävä

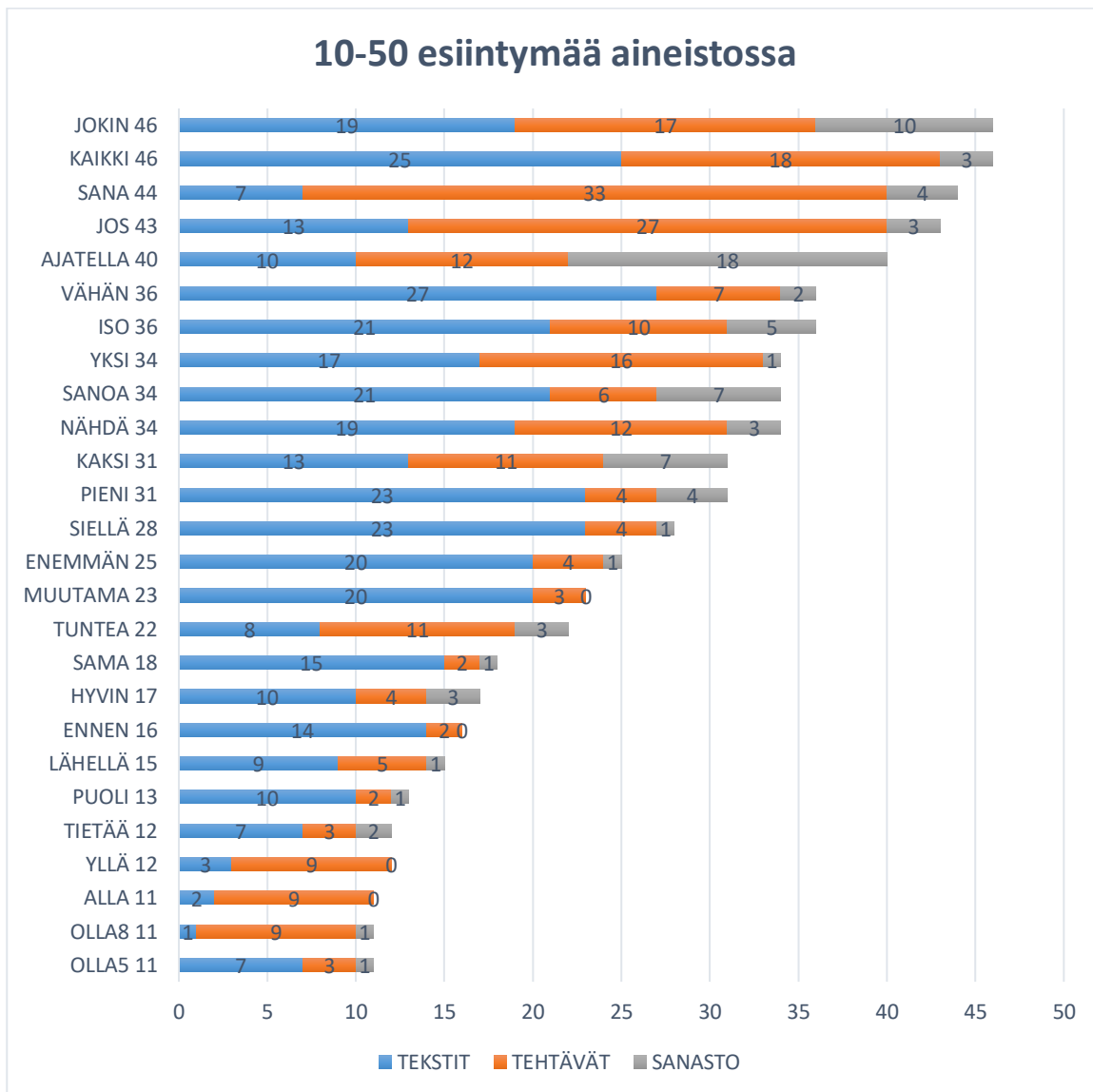
sana Suomen kielen taajuussanastossa on *sellainen*, joka on sijalla 61. Myös sanat *jälkeen*, *tässä* ja *koska* ovat aineistossani korkealla, vaikka Suomen kielen taajuussanastossa ne ovat vasta samassa järjestyksessä sijoilla 70, 75 ja 93. Sanat *missä* ja *tahtoa* eivät puolestaan kuulu em. mukaan suomen kielen sadan yleisimmän sanan joukkoon lainkaan. Muuten diagrammin sanat vastaavat ihan hyvin Suomen kielen taajuussanaston listausta.

Sana *paljon* on huomattavasti yleisempi kuin sen negaatio *vähän*: Sanalla *paljon* (tai *moni*) on aineistossa esiintymiä 65, kun *vähän* (tai *harva*) esiintyy aineistossa vain 36 kertaa. Myös paikanilmaus *tässä* (tai *täällä*) on huomattavasti yleisempi kuin ilmaukset *tuolla* (tai *tuossa*) tai *siellä* (tai *siinä*). *Tässä* esiintyy aineistossa 63 kertaa, kun sanalla *siellä* on aineistossa 28 esiintymää ja sanalla *tuolla* vain 4.

Sana *ennen* yleisenä menneisyyttä kuvaavana ilmauksena sekä sana *ennen* johonkin konkreettiseen ankkuroituna ilmauksena ovat jokseenkin harvinaisia: Sanalla *ennen* (yleinen) on aineistossa vain 16 esiintymää ja sanalla *ennen* (jotakin) on vain 4. *Jälkeen* puolestaan on huomattavan yleinen: Sillä on esiintymiä peräti 60.

Myös tämä diagrammi tukee sitä havaintoa, että alkusanat ovat määränsä puolesta selkeästi enemmän läsnä teksteissä ja tehtävissä kuin sanasto-osioissa. Tähän tekee poikkeuksen alkusana -LAINEN/-LÄINEN: Sen jokseenkin vahvaa edustusta sanastollisilla osa-alueilla selittää ensisijaisesti kappaleen 1 ”Minä olen suomalainen/ruotsalainen/kenialainen” -tyyppiset osiot.

Diagrammi 3 – alkusanojen määrä aineistossa 3/5

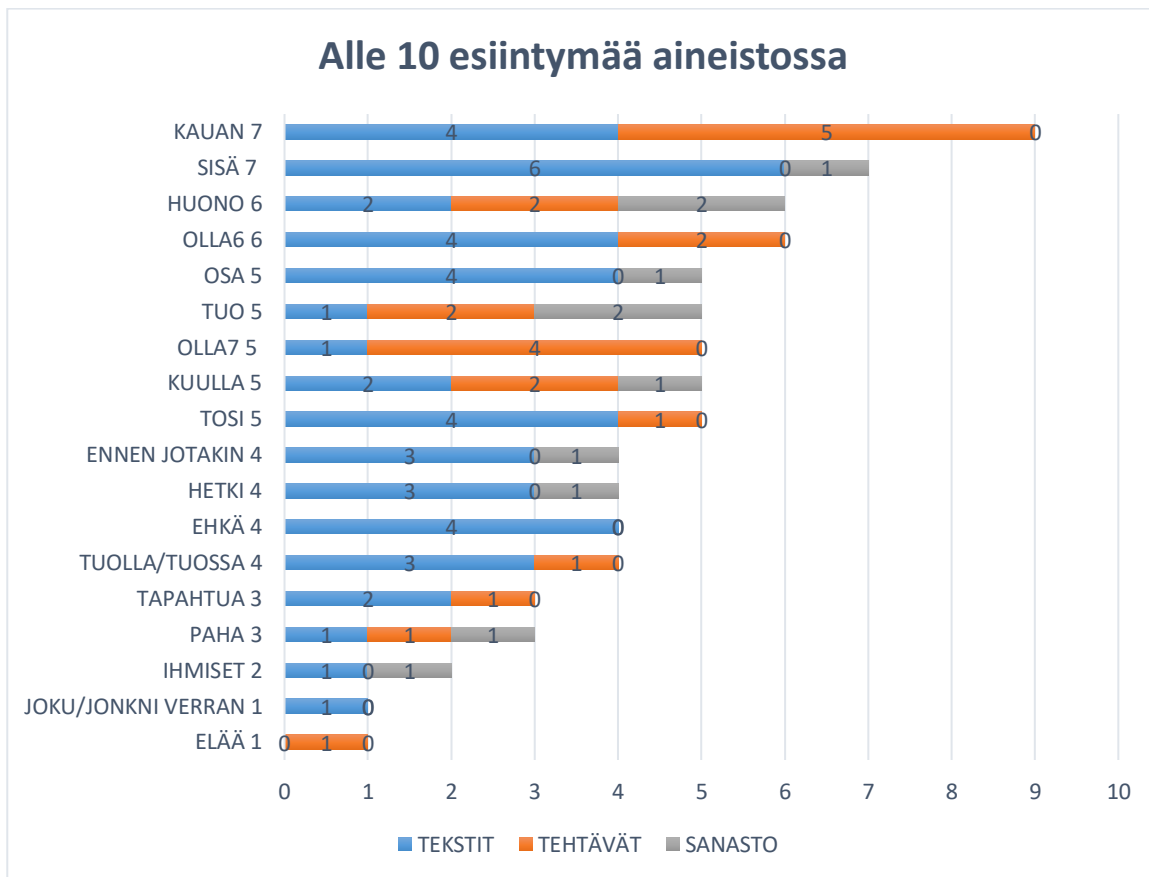


Tämä diagrammi, joka siis esittää sellaisten sanojen määrän aineistossa, joita on yli 10 mutta alle 50 on diagrammeista suurin. Toisin sanoen näitä sanoja on enemmän kuin mihinkään muuhun diagrammiin kuuluvia sanoja. Diagrammissa ei ole mainittavia eroja Suomen kielen taajuussanastoon (1979) verraten.

Sanat *iso* ja *pieni* kuuluvat molemmat tähän kaavioon, kuten myös sanat *yksi* ja *kaksi*. Niillä ei ole paljoakaan eroja määrässä.

Kuten edellisetkin diagrammit, myös tämä tukee havaintoa, jonka mukaan sanasto-osiot sisältävät vähemmän alkusanoja kokonaismäärällisesti. Tämän diagrammin sanoissa liikutaan jo niin pienten määrien keskuudessa, että mukana on jo alkusanoja, joita ei ole sanasto-osioissa lainkaan: MUUTAMA, ENNEN, YLLÄ ja ALLA.

Diagrammi 4 – alkusanojen määrä aineistossa 4/5



Tämä diagrammi sisältää sanat, joita on aineistossa alle kymmenen, mutta enemmän kuin nolla. Kokonaisuudessaan katson, että näiden sanojen esiintymät ovat jo niin harvinaisia, että niistä on hankala tulkita paljoakaan. Kuten saksankielinen sanonta kuuluu, *einmal ist keinmal* – kerran ei ole kertaakaan.

Mielenkiintoinen huomio on sanan *ihmiset* vähyys: sanalla on vain 2 esiintymää. Syy, miksi katson tämän olevan mielenkiintoista, on se, että kuten kirjoitin luvussa 5, englanninkielisessä NSM-paradigmassa sana *people* katsotaan alkusanaksi sellaisenaan monikkomuodossa. Jos puolestaan verrataan sanaa IHMISET sanaan HÄN (jonka aloleksi on IHMINEN), huomataan eron olevan erittäin suuri: Sanalla HÄN on esiintymiä aineistossa jopa 223.

Mitä pienemmissä määrissä liikutaan alkusanojen kokonaismäärän suhteen, sitä pienemmäksi käy tilastollinen evidenssi eri alkusanojen jakautumisesta kullekin osa-alueelle. KAUAN, OLLA6, OLLA7, TOSI, EHKÄ, TUOLLA, TAPAHTUA, JOKU ja ELÄÄ eivät saa yhtään esiintymiä sanastollisissa osa-alueissa. Mielenkiintoisena noottina voidaan mainita, että mikään alkusana kirjassa ei esiinny vain sanasto-osioissa.

Diagrammi 5 – alkusanojen määrä aineistossa 5/5



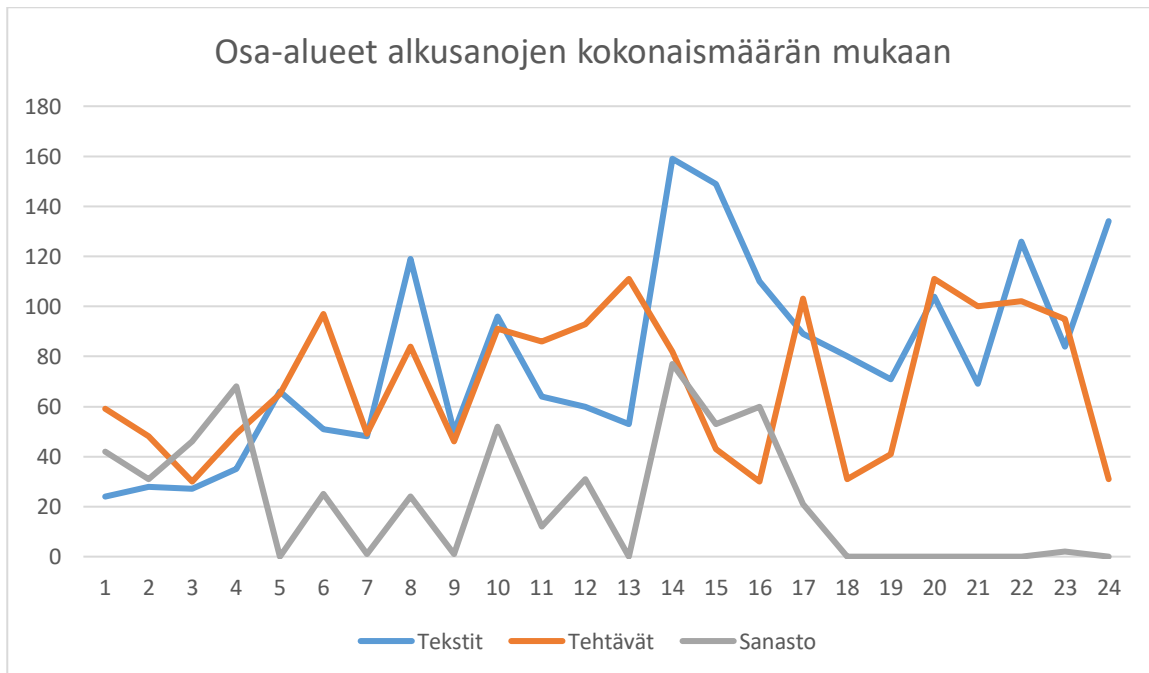
Tämä diagrammi kuvaa niitä sanoja, jotka eivät esiinny aineistossa kertaakaan. Syitä tälle voidaan vain spekuloida, joskin sellaisen kuin *kuolla* puuttuminen aineistosta ei liene yllätys. Sen sijaan olisin olettanut, että *kaukana*, *koskea* tai *liikkua* olisivat esiintyneet aineistossani edes kerran. On tosin hyvä huomata, että *liikkua* on melko tekninen sana, eikä se sellaisenaan oman kielikorvani mukaan esiinny kielessä kovinkaan paljoa. Sen sijaan liikkumista kuvataan sellaisilla sanoilla kuin *mennä* ja *tulla*. Myös sanan *ruumis* olisin olettanut esiintyvän aineistossani edes synonyymisenä esiintymänä *keho*, *vartalo* tai *kroppa*. Ilmaisista *jonkin aikaa* ja *vähän aikaa* kirjoitin jo luvussa 5. Olen kuitenkin hieman yllättynyt, että ne eivät esiintyneet aineistossa kertaakaan.

Ensimmäinen näistä sanoista Suomen kielen taajuussanastossa (Saukkonen ym. 1979) on KOSKEA, joka on sijalla 149. Tämän jälkeen tulee LIIKKUA sijalla 580, KUOLLA sijalla 682, RUUMIS sijalla 1197 ja KAUKANA sijalla 1798. JONKIN AIKAA ja VÄHÄN AIKAA eivät kahden sanan rakenteina teoksesta löydy.

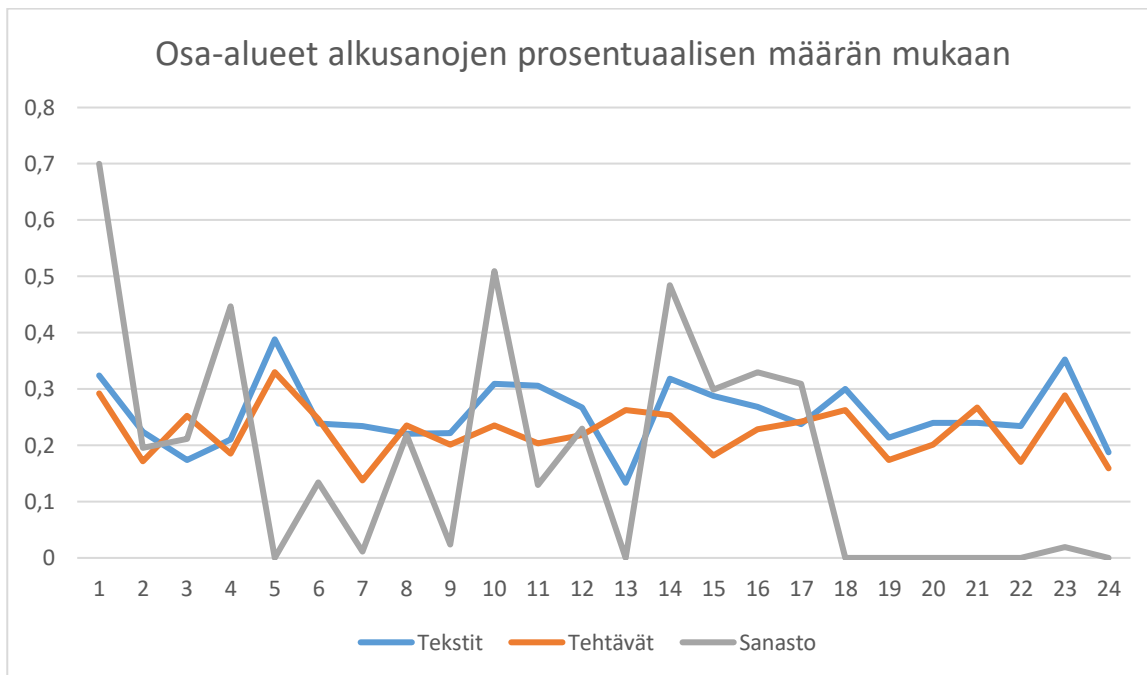
6.2 Osa-alueiden välinen vertailu

Tässä alaluvussa pyrin vastaamaan kysymykseen siitä, korostuuko jokin kirjan osa-alueista erityisen paljon alkusanojen määrän ja suhteellisen määrän mukaan. Havainnollistan kysymystä kahdella diagrammilla, joissa näytetään kunkin osa-alueen alkusanat rinnan. Ensimmäisessä diagrammissa kyseessä on kokonaissanamäärä, ja toisessa prosentuaalinen määrä.

Diagrammi 6 – alkusanojen kokonaismäärä osa-alueittain



Diagrammi 7 – alkusanojen suhteellinen määrä osa-alueittain



Ensimmäisestä diagrammista on nähtävissä, että tekstialueet sisältävät jonkin verran enemmän alkusanoja kuin tehtävät. Sanastokokonaisuudet puolestaan sisältävät niitä melko vähän. On syytä mainita, että sanastokohdan diagrammin nollakohdat merkitsevät kappaleita, joissa kyseisiä osa-alueita ei edes ole.

Jälkimmäisen diagrammin merkitys korostuu, kun verrataan alkusanojen määrää suhteessa osa-alueiden sanamääriin: Tekstien sisältämien alkusanamäärien voidaan katsoa olevan myös suhteellisesti korkeammat kuin tehtävien, joskaan ei kovin paljoa. Vain kappaleissa 3, 13 ja 21 tehtävät sisältävät enemmän alkusanoja prosentuaalisesti kuin tekstit. Suurin ero tekstien ja tehtävien välillä vaikuttaa olevan kappaleissa 7 (tekstit: 23,4%, tehtävät: 13,8%), 10 (tekstit: 30,9%, tehtävät: 23,5%), 11 (tekstit: 31,1%, tehtävät: 20,4%) ja 15 (tekstit: 28,8%, tehtävät: 18,1%). Tekstiosioiden alkusanojen prosentuaalinen vaihteluväli on 13,3% (kappale 5) – 38,8% (kappale 13). Tehtävien alkusanojen prosentuaalinen vaihteluväli puolestaan on 13,8% (kappale 7) – 33% (kappale 5).

Sen sijaan sanastollisista kokonaisuuksista on hankalampi löytää johdonmukaisuutta. Kuvaajan vaihtelu on suurta: Ensimmäisen kappaleen sanastollisissa osa-alueissa alkusanojen määrä on jopa 70%, kun taas pienin suhteellinen alkusanojen määrä sellaisessa kappaleessa, jossa sanastollisia osa-alueita on, on kappaleessa 7; 1,1%. Mielestäni kuitenkin sanastollisten alueiden poukkoilevaa kuvaajaa selittää esimerkiksi se asia, että sanastollisten osa-alueiden sanamäärä on huomattavasti pienempi kuin kahden muun: Yksikin kimppu alkusanoja saattaa nostaa myös niiden

prosentuaalisen määrän kattoon, kun taas pieneen osa-alueeseen, kuten kappaleen 7 88:aan sanaan ei välttämättä satu montaakaan alkusanaa.

Vastauksena kysymykseen ”ovatko alkusanat enemmän edustettuina jollain tietyllä osa-alueella kuin muilla?” voin siis vastata, että tekstikokonaisuudet tosiaan sisältävät enemmän alkusanoja kuin tehtäväkokonaisuudet. Samaa ei voi luotettavasti kuitenkaan todeta sanastollisista osa-alueista niiden pienen sanamäärän ja sen myötä epäjohdonmukaisen kuvaajan takia.

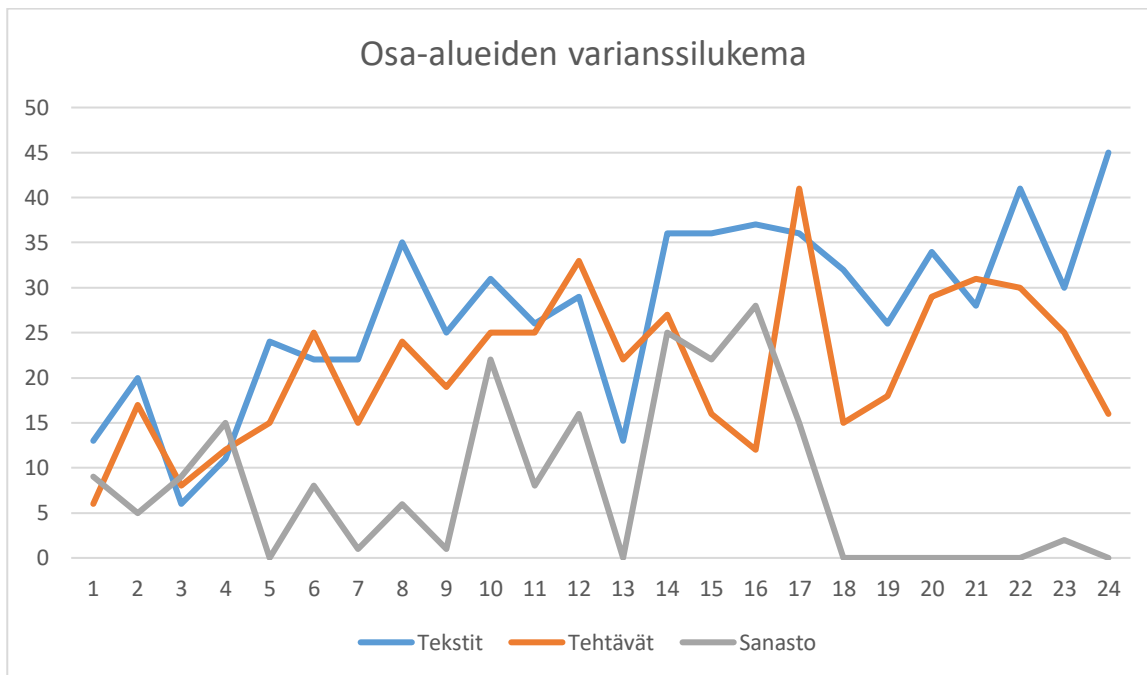
Alkusanojen kokonaismäärä ja prosentuaalinen määrä eivät kuitenkaan tarjoa kaikkea tilastollisesti merkittävää dataa. Esimerkkinä tilanteesta, jossa sekä alkusanojen prosentuaalinen määrä että kokonaismäärä ovat korkeat, mutta jossa alkusanat ovat suppeasti edustettuina olisi esimerkiksi tehtävä, joka koostuisi lähes kokonaan tietyn persoonapronominin, esimerkiksi sanan *minä* esiintymistä, joiden perusteella tulisi taivuttaa tiettyjä verbejä. Tällainen tehtävä sijoittuisi siis edellisissä kaavioissa mahdollisesti korkealle, mutta sen arvo alkusanojen suhteen olisi kuitenkin heikko.

Tällaisia tapauksia varten haluan vielä erikseen ottaa mukaan käsitteen, jota kutsun alkusanojen varianssiksi. Tämä tarkoittaa sitä, miten moni alkusana on edustettuna missäkin kappaleessa tai osa-alueessa. Luku voi siis periaatteessa olla mitä tahansa nollan ja kaikkien aineistooni hyväksymieni alkusanojen määrän eli 76:n välillä, mutta koska 7 alkusanaa eivät esiinny aineistossa lainkaan, se ei oikeasti voi nousta kuin 69:ään, ja tämäkin on lähinnä teoreettista: mikään osa-alue ei yllä näin korkeaan lukemaan.

Jotta siis saadaan vielä yksi oleellinen faktori eri osa-alueiden välille, tulee kunkin osa-alueen varianssiarvo laskea kaavalla $76-x$, jossa x kuvaa niitä alkusanoja, joiden määrä kyseisellä osa-alueella on 0. Jos jokin osa-alue siis sisältää vain 10 nolla-arvoa, sen varianssiluku on 66, joka on erittäin korkea, sillä se tarkoittaa, että kyseinen osa-alue sisältää esiintymiä 66:sta eri alkusanasta.

Olen laskenut osa-alueiden varianssiarvot kappaleittain, ja esitän ne seuraavassa kaaviossa:

Diagrammi 8 – osa-alueiden varianssiarvo



Kaaviosta on nähtävillä se, että prosentuaalisesti suuremaan määrän lisäksi tekstiosa-alueilla on keskimäärin suurempi varianssi kuin tehtävissä. Sanasto-osa-alueet puolestaan ovat keskimäärin matalammalla varianssin suhteen. Tämä tarkoittaa siis sitä, että sanastolliset osa-alueet sisältävät alkusanoja pienellä varianssilla, toisin sanoen ne siis sisältävät samoja alkusanoja, jotka toistuvat. Varianssikäyrä on myös nouseva sekä tekstien että tehtävien osalta, mistä lisää seuraavassa alaluvussa.

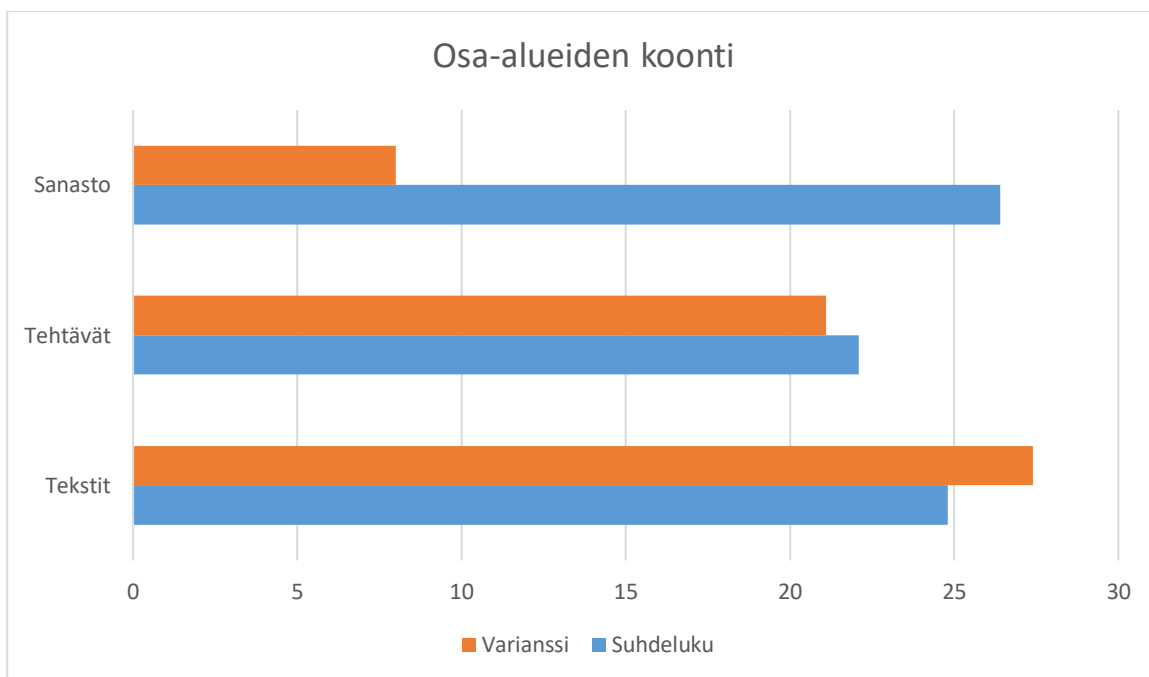
Vaikka teksteissä on keskimäärin suurempi varianssi, voidaan todeta tehtävien olevan koko kirjan matkalta melko lähellä niitä. Kappaleissa 3, 6, 12, 13, 17 ja 21 tehtäväosioiden varianssi on itseisassa jopa suurempaa kuin teksteissä. Suurimmat erot tehtävien ja tekstien välillä puolestaan on nähtävissä kappaleissa 15 (tekstit: 36, tehtävät: 16), 16 (tekstit: 37, tehtävät: 12), 18 (tekstit: 32, tehtävät: 15) ja 24 (tekstit: 45, tehtävät: 16).

Lisäksi voidaan verrata myös koko osa-alueiden varianssiarvojen keskiarvoja ja mediaaniarvoja: Keskiarvo kertoo osa-alueista yleisesti vertailtavan arvon, ja mediaani esittää koko arvokoukon keskimmäisen arvon. Tekstien varianssilukujen keskiarvo on 27,4 ja mediaani 28,5. Tehtävien varianssilukujen keskiarvo on 21,1 ja mediaani 20,5. Sanasto-osioiden varianssilukujen keskiarvo on 8 ja mediaani 5,5.

Osa-alueiden varianssiarvoista on nähtävillä siis se, miten tasaisesti alkusanat esiintyvät milläkin osa-alueella. Se ei itsessään kerro paljoakaan siitä, kuinka hyvin alkusanat ovat millä-

kin kirjan osa-alueella edustettuina, mutta se tarjoaa oleellisen vastapainon alkusanojen kokonaismäärälle sekä prosentuaaliselle määrälle: Jos alkusanojen kokonaismäärä ja prosentuaalinen määrä on osa-alueella korkea ja varianssiluku on suuri, voidaan olettaa osa-alueen sisältävän alkusanoja jokseenkin kattavasti. Kootakseni tämän luvun tiedon esitän vielä pylväsdiagrammin, jossa havainnollistan osa-alueiden suhteellisen alkusanamäärän alkusanojen varianssilukujen keskiarvojen rinnalla:

Diagrammi 9 – osa-alueiden koonti



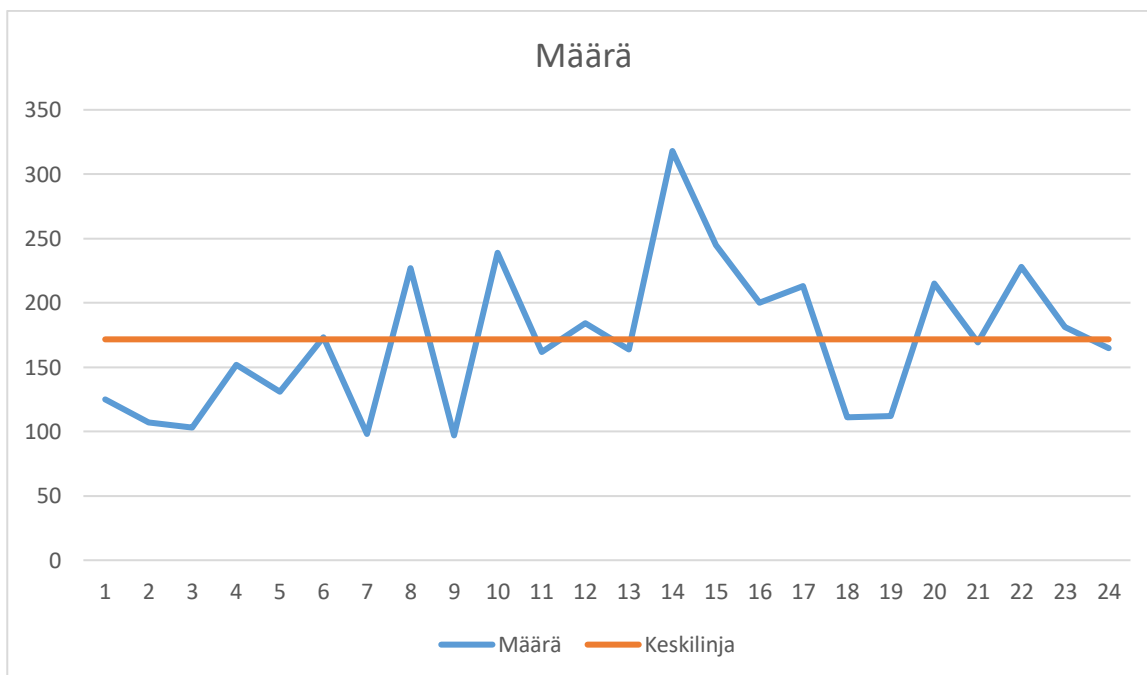
Tämä diagrammi tukee sitä huomiota, että alkusanat ovat parhaiten edustettuna kirjan tekstiosioissa. Ne ovat verrattain hyvin edustettuina myös tehtäväosa-alueilla, mutta eivät aivan niin hyvin kuin teksteissä. Sanasto-osa-alueilla puolestaan on suhteellisesti eniten alkusanoja, mutta niiden varianssi on pieni. Tämä tarkoittaa sitä, että vaikka alkusanoja olisi paljon, ne edustavat vain pientä määrää kaikista alkusanoista.

6.3 Kappaleiden välinen vertailu

Tässä luvussa pyrin tarkastelemaan sitä, onko kirjan eri kappaleiden välillä eroja alkusanojen kokonaismäärässä, suhteellisessa määrässä ja varianssissa. Tämän perusteella voidaan muodostaa päätelmiä siitä, missä vaiheessa alkusanoja eniten tuodaan oppijoiden syötteeseen: Aloitetaanko niiden tuominen heti kirjan alussa, lisääntyvätkö ne kirjan edetessä vai ovatko ne tasaisesti läsnä?

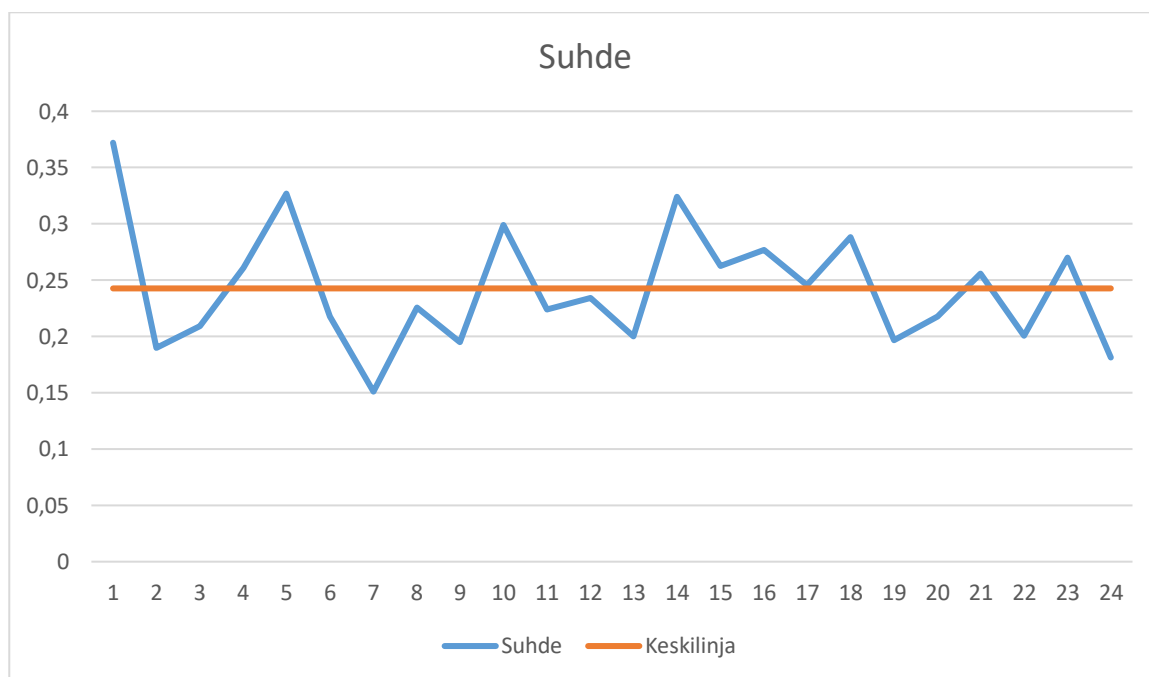
Tarkastellakseni alkusanojen kokonaismääriä, suhteellisia määriä ja varianssia kussakin kappaleessa esitän myös näitä asioita kuvaavia kaavioita. On kuteinkin syytä mainita, että kaaviossa, jossa esitelen kappaleiden keskimääräisen varianssin, otan erikseen huomioon sen, että kaikissa kappaleissa ei ole sanastoa opettavia osa-alueita. Tämä tarkoittaa sitä, että näiltä osin en voi myöskään laskea tällaisten kappaleiden osalta varianssin keskiarvoa kaikista kolmesta kirjan osa-alueesta, sillä varianssin arvo 0 vaikuttaisi oleellisesti keskiarvolukemaan. Käytännössä: Kappaleessa 24 on tekstiosa-alueessa 134 alkusanaa ja tehtävissä 31. Sanasto-osa-alueita ei kappaleessa ole. Tekstien varianssi puolestaan on 45 ja tehtävien 16. Tämä tarkoittaa sitä, että kappaleen varianssi on 30,5; kappaleessa on siis keskimäärin edustettuina 30,5 alkusanaa.

Diagrammi 10 – alkusanojen kokonaismäärä kappaleittain



Kaaviosta on nähtävillä, että alkusanojen kokonaismäärän vaihtelu on jokseenkin suurta. Vaihteluväli on 97 (kappale 9) – 318 (kappale 14). Tätä selittää kylläkin vahvasti kappaleiden ja niiden sisältämien tekstien, tehtävien ja sanastollisten kokonaisuuksien pituus ja määrä. Tilastollisesti merkitsevät, eli ympäristöstään poikkeavat kappaleet ovat kuitenkin 1, 2, 3, 7, 8, 9, 10, 14, 18, 19, 20 ja 22. Kappaleissa 1, 2, 3, 7, 9, 18 ja 19 alkusanoja on erityisen vähän ja kappaleissa 8, 10, 14, 20 ja 22 niitä on erityisen paljon. Näitä kappaleita on syytä tarkastella myös alkusanojen suhteellista määrää kuvaavan kaavion valossa.

Diagrammi 11 – alkusanojen suhteellinen määrä kappaleittain



Myös suhteellinen vaihtelu on jokseenkin suurta; vaihteluväli on 15,1% (kappale 7) – 37,2% (kappale 1). Ympäristönsä puolesta poikkeavia kappaleita ovat erityisesti 1, 5, 7 ja 14. Tästä voidaan päätellä parikin asiaa: Ensinnäkin poikkeavien kappaleiden vähyys suhteellisesti osoittaa edellisessä diagrammissa näkyvän suuren vaihtelun selittyvän hyvinkin pitkälti kappaleiden sanamäärällä. Vain kappaleet 1, 7 ja 14 ovat keskilinjasta poikkeavia molemmissa kaavioissa.

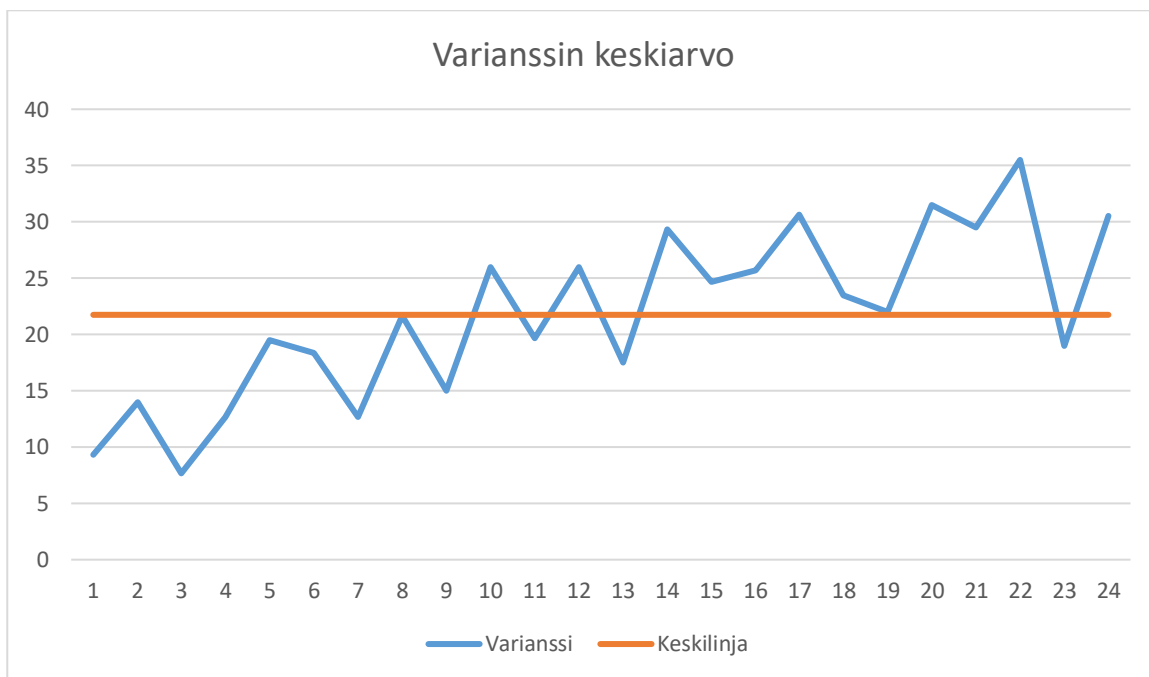
Kappaleessa 1 on siis suhteellisesti huomattavan paljon alkusanoja. Kokonaisuudessaan niitä puolestaan on keskiarvoa selkeästi vähemmän. Tämä tarkoittaa sitä, että vaikka kappaleessa ei ole paljon tekstiä, se koostuu erityisen suurelta osin alkusanoista.

Kappaleessa 7 puolestaan on sekä kokonaismäärältä että suhteelliselta määrältä vähän alkusanoja. Tämä tarkoittaa sitä, että kappaleessa on kohtalaisen paljon sanoja ylipäättään, mutta vain pieni osa niistä on alkusanoja.

Kappaleessa 14 alkusanoja on paljon sekä kokonaismäärältä että suhteelliselta määrältä. Tämä tarkoittaa sitä, että niitä pakkautunut erityisen paljon kyseiseen kappaleeseen.

Näistä huomioista on vieläkin pääteltävissä jokseenkin vähän, joten on syytä ottaa mukaan vielä kolmas kaavio, joka kuvaa kappaleiden keskimääräistä varianssia alkusanojen suhteen.

Diagrammi 12 – alkusanojen keskimääräinen varianssi kappaleittain



Tästä taulukosta näkee sen, että toisin kuin alkusanojen määrä ja suhteellinen määrä, niiden varianssi lisääntyy sitä mukaa, mitä pidemmälle kirjassa edetään. Vasta kappaleessa 10 varianssi nousee lähes pysyvästi paria notkahdusta lukuun ottamatta yli keskilinjan, jonka arvo on 21,743. Puolestaan ennen kappaletta 10 alkusanojen määrä pysyttelee yhtä poikkeusta lukuun ottamatta tämän linjan alapuolella. Tämä on kiintoisa huomio, sillä se osoittaa alkusanojen määrän lisääntyvät syötteessä sitä mukaa, mitä pidemmälle kirjassa edetään.

Kun tarkastellaan kappaleita 1, 7 ja 14, jotka edellisissä kaavioissa korostuivat ympäristöönsä nähden, voidaan myös päätellä jotain niiden luonteesta. Kappaleessa yksi, jossa siis on suhteellisesti paljon mutta määrällisesti vähän alkusanoja, varianssiarvo on vain 9,33. Tämä tarkoittaa sitä, että kourallinen alkusanoja toistuu kappaleessa useasti.

Kappaleessa seitsemän puolestaan, jossa siis on alkusanoja vähän sekä kokonaismäärällisesti että suhteellisesti, on myös pieni varianssiluku: 12,667. Tämä tarkoittaa sitä, että kyseinen kappale sisältää vähän alkusanoja kaikilla mittareilla.

Kappaleessa 14, jossa on siis sekä määrällisesti että suhteellisesti paljon alkusanoja, on myös ympäristöönsä nähden huomattavan korkea varianssi: 29,33. Tämä tarkoittaa sitä, että kyseisessä kappaleessa on paljon alkusanoja kaikilla mittareilla. On myös huomionarvoista, että kappaleen 14 jälkeen alkusanojen keskimääräinen varianssi pysyttelee selkeästi korkeammalla kuin ennen kappaletta 14. Sitä voidaan pitää siis jonkinlaisena käännoispisteenä kirjassa.

6.4 Alkusanakohtainen jakauma

Tässä luvussa esitän tarkemmin kaavioilla, miten mikäkin alkusana on edustettuna aineistossa. Esitän alkusanat samoissa kategorioissa kuin missä olen ne luvussa 5 esitellyt. Tällöin niistä voi tarkastella myös sitä, esiintyvätkö samaan kategoriaan kuuluvat alkusanat kirjassa eri tavalla.

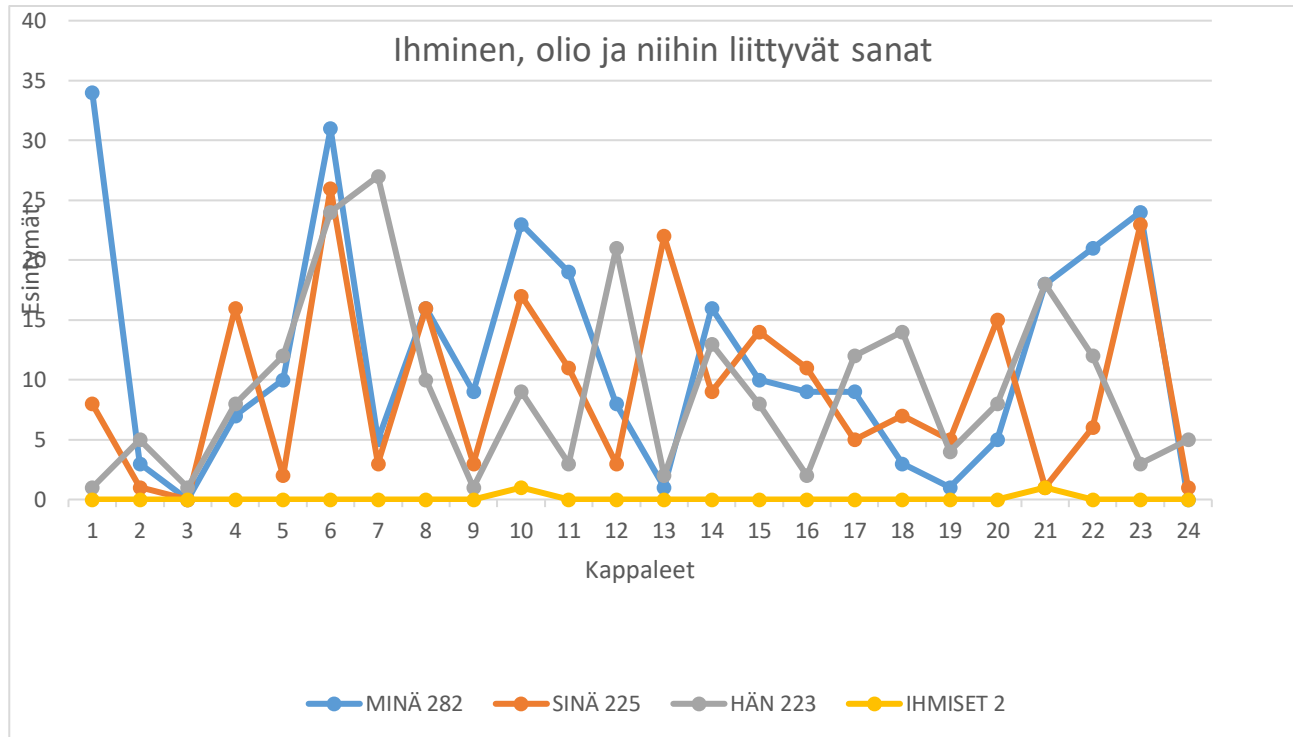
Diagrammien avulla voin tarkastella, ovatko alkusanat jakautuneet aineistoon tasaisesti, lisääntykö tai vähentykö niiden määrä kirjan loppua kohden ja onko niiden määrissä esimerkiksi piikkejä tai kuoppia. Erityisesti sellaisten piikkien kohdalla, jotka poikkeavat selkeästi kyseisen sanan kuvaajasta, voin tarkentaa havaintojani kyseisestä alkusanasta tarkastelemassa sitä kirjaympäristössä. Esimerkiksi sana, joka ei ole kovin paljon edustettuna aineistossa, mutta jolla on yksi selkä piikki, on kiintoisa siksi, että tämä piikki todennäköisesti selittää sen esiintymistä aineistossa. Samaan tapaan sanan, jolla on useampia piikkejä tai ei piikkejä lainkaan, voidaan katsoa esiintyvän aineistossa jokseenkin tasaisesti. Luvun diagrammeista näkyy myös kyseisen diagrammin alkusanojen esiintymien määrä koko aineistossa sanan perässä, esim. MINÄ 282.

Diagrammien jälkeen esittelen taulukon, joka kuvaa alkusanoja tietyin numeerisin arvoin. Ensimmäinen arvo on alkusanan kokonaismäärä aineistossa. Toinen on luku, jota kutsun alkusanan jakaumaluvuksi. Tämä tarkoittaa siis sitä, kuinka monessa kappaleessa kyseinen alkusana esiintyy. Jakamalla alkusanan kokonaismäärä aineistossa sen jakaumaluvulla saadaan alkusanan keskimääräinen määrä kappaleissa, joissa se esiintyy.

En ole tässä luvussa ottanut lainkaan huomioon niitä seitsemää alkusanaa, jotka eivät esiinny aineistossani kertaakaan. Niillä ei ole tilastollisesti enempää arvoa kuin tieto siitä, että niitä ei aineistossa ole.

6.4.1 Ihminen, oli ja niihin liittyvät sanat

Diagrammi 13 – ihminen, olio ja niihin liittyvät sanat



Diagrammista on nähtävissä, että MINÄ on kaikista tämän kategorian sanoista yleisin kahteen otteeseen; kappaleissa 1 ja 6 sen määrät ovat suuremmat, kuin millään muulla kategorian alkusanalla missään kirjan kappaleessa. Alkusanalla MINÄ on havaittavissa 5 selkeää piikkiä, kappaleissa 1, 6, 10, 14 ja 23. Alkusanalla SINÄ piikkejä puolestaan on 8, kappaleissa 1, 4, 6, 8, 10, 13, 20 ja 23. Alkusanalla HÄN piikkejä on 6, kappaleissa 7, 10, 14, 18 ja 21. Sana IHMISET on niin harvinainen, että piikeistä ei voi puhua.

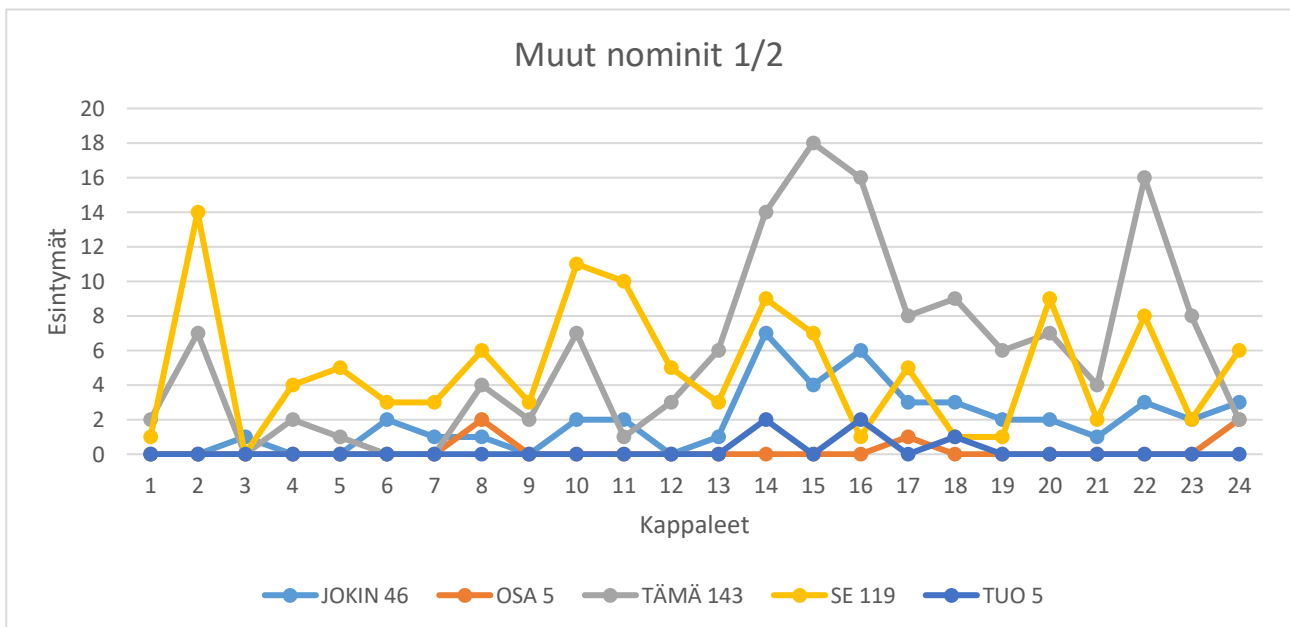
Piikkien lukumäärästä voi päätellä, että vaikka MINÄ on aineistossa jonkin verran yleisempi (282) kuin SINÄ (225) ja HÄN(223), esiintyvät kaksi viimeksi mainittua kirjassa tasaisemmin. Alkusanan MINÄ jakaumaluku on 22, sanan SINÄ 23 ja sanan HÄN 24. Tämä tarkoittaa sitä, että alkusanan MINÄ edustuksia on keskimäärin kappaletta kohden 12,82, sanan SINÄ 9,78 ja sanan HÄN 9,29.

Taulukko 2 – ihminen, olio ja niihin liittyvät sanat

	Määrä	Jakauma	Määrä/jakauma
MINÄ	282	22	12,82
SINÄ	225	23	9,78
HÄN	223	24	9,29
IHMISET	2	2	1

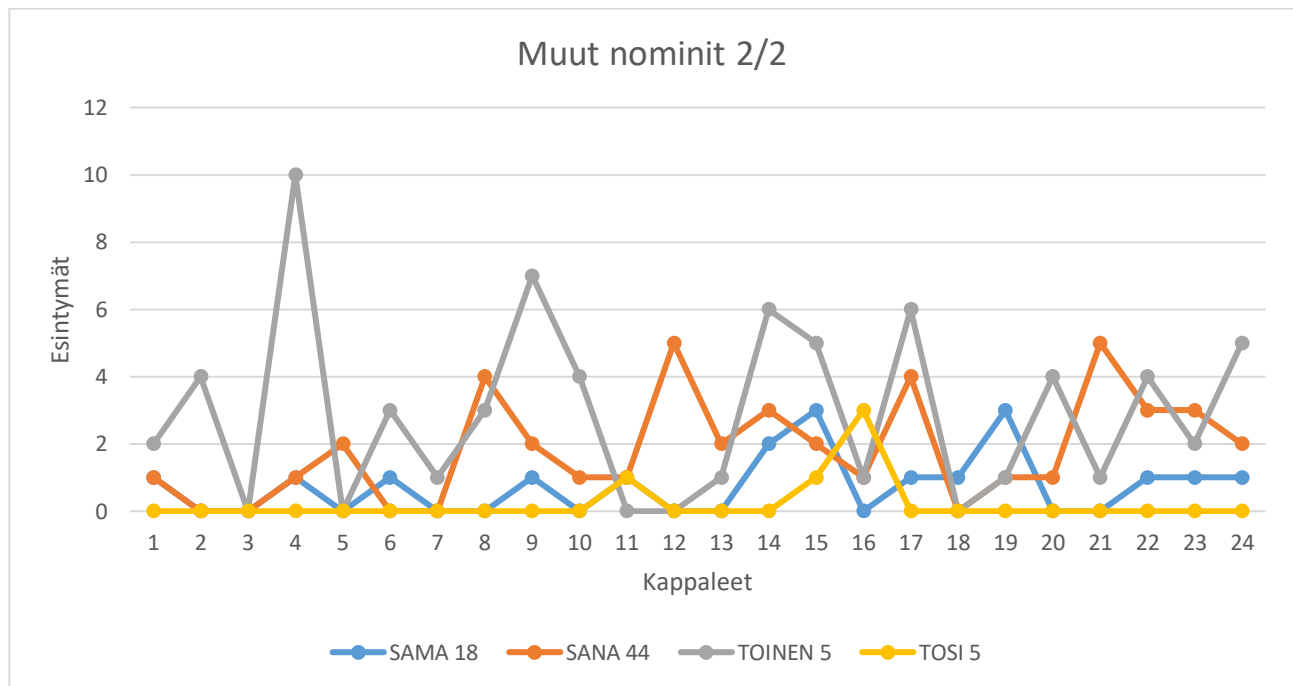
6.4.2 Muut nominit

Diagrammi 14 – muut nominit 1/2



Diagrammeista on nähtävillä, että muutoin tasainen TÄMÄ kirjan loppupuolen kappaleissa 14, 15, 16 ja 22 huomattavasti. Sanalla SE puolestaan on huomattava piikki kappaleessa 2, ja myöhemmin kirjassa sillä on pienempiä piikkejä. Sekin vaikuttaa sikäli esiintyvän kirjassa jokseenkin tasaisesti. Sanat TUO ja OSA ovat kirjassa harvinaisia, kuten myös JOKIN kirjan alkupuolella, joskin sille kertyy esiintymiä erityisesti kappaleissa 14, 15 ja 16. TÄMÄ ja SE esiintyvät kirjassa kuitenkin hyvin yleisesti (21, 23), samoin kuin JOKIN (18). Sanojen OSA ja TUO pieni jakaumaluku selittyy niiden pienellä kokonaismäärällä.

Diagrammi 15 – muut nominit 2/2



Alkusanaa TOINEN ei voine kutsua tasaisesti esiintyväksi: Sillä on suhteellisen suuria piikkejä neljä, ja muulloin se on jokseenkin harvinainen. Sanalla TOSI on myös yksi selkeä piikki, joka ei tosin ole järin merkittävä sanan harvinaisuuden vuoksi: Kappaleessa 16 sillä on 3 esiintymää, mutta tämä ei vielä ole tilastollisesti merkittävää. SAMA esiintyy kirjassa jokseenkin tasaisesti, sanalla SANA on jokunen pienehkö piikki, mutta ne eivät myöskään vaikuta tilastollisesti merkittäviltä.

Jälkimmäisen diagrammin sanoista suurin jakaumaluku on sanalla SANA (19), myös SAMA (13) esiintyy jokseenkin yleisenä. TOINEN ja TOSI ovat molemmat hyvin harvinaisia.

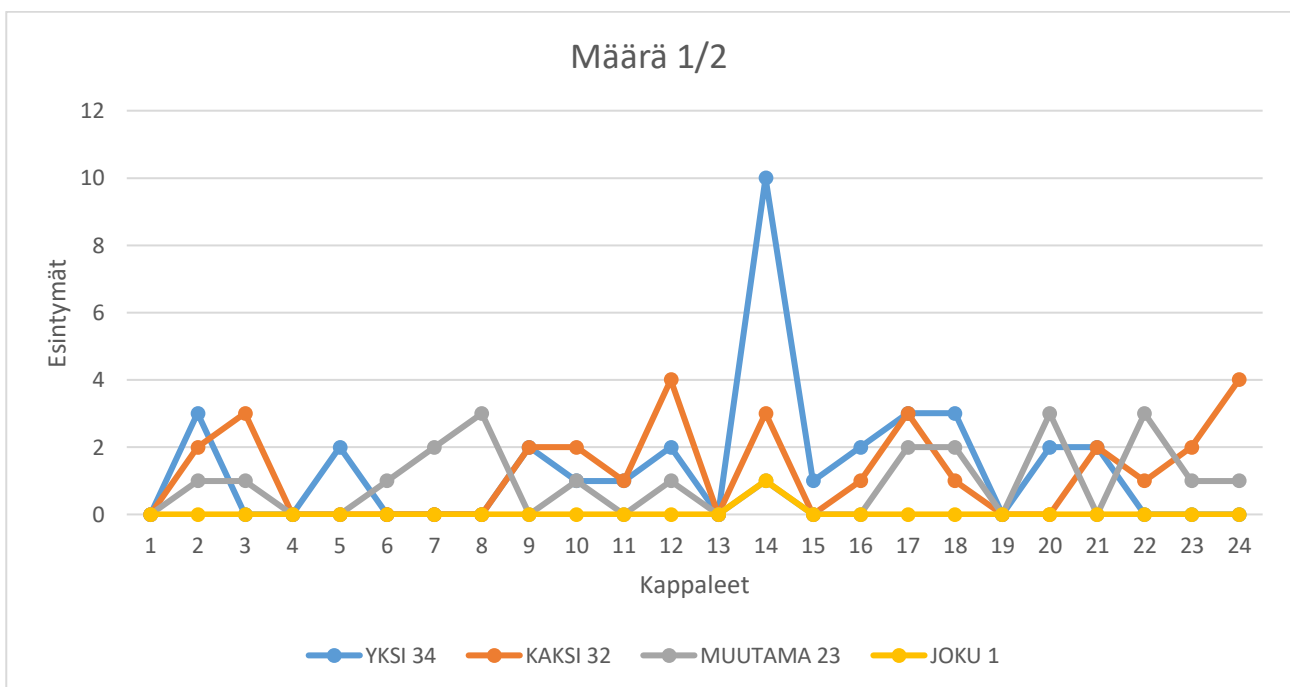
Taulukko 3 – muut nominit

	Määrä	Jakauma	Määrä/jakauma
JOKIN	46	18	2,56
OSA	5	3	1,667
TÄMÄ	143	21	6,81
SE	119	23	5,17
TUO	5	3	1,667

SAMA	18	13	1,38
SANA	44	19	2,32
TOINEN	5	3	1,667
TOSI	5	3	1,667

6.4.3 Määrä

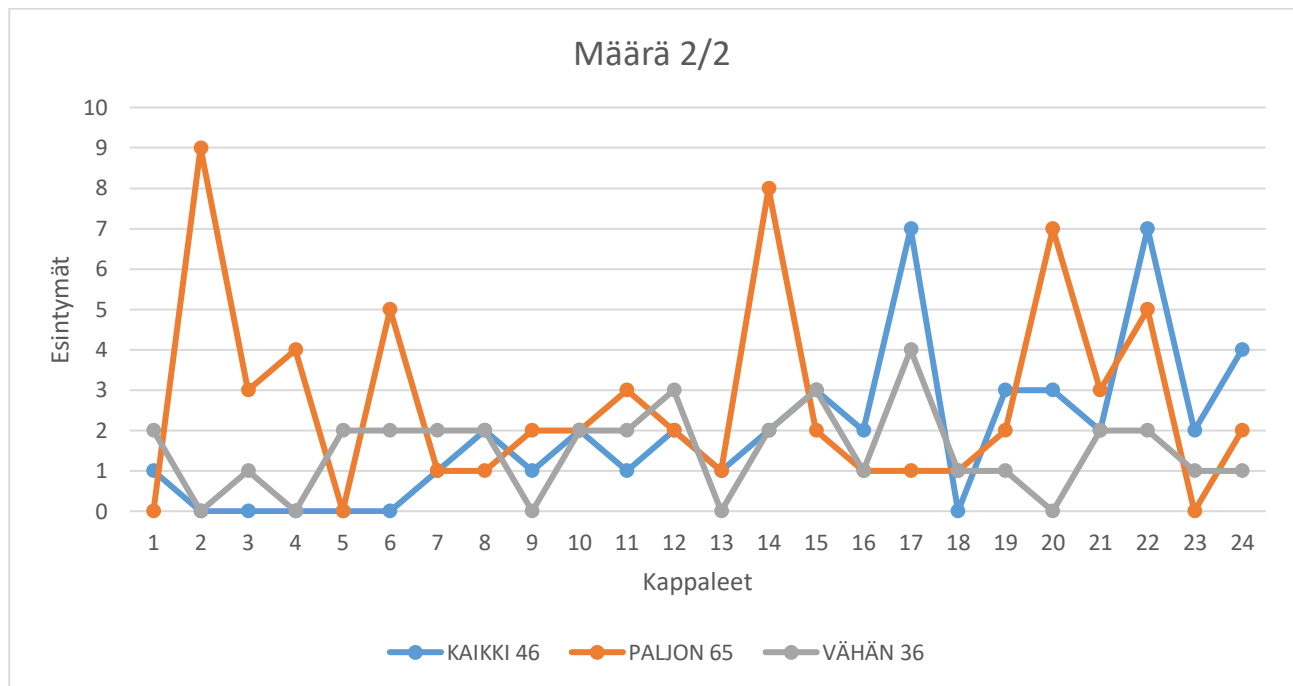
Diagrammi 16 – määrä 1/2



Ainoa selkeä piikki tässä diagrammissa on alkusanan YKSI 10 esiintymää kappaleessa 14. Kyseisen kappaleen tehtäväosiossa on tehtävä, jossa opetellaan ilmaisua ”yksi jollekin, ja yksi jollekin toiselle”. Tämä selittää piikin suurelta osin. Sanat KAKSI ja MUUTAMA esiintyvät aineistossa jokseenkin tasaisesti. Sana JOKU puolestaan esiintyy aineistossa tasan kerran, joten sitä ei tohtine piikkiä kutsua.

YKSI on läsnä 13:ssa kirjan kappaleessa, KAKSI ja MUUTAMA 14:ssä. Sanat ovat siis jakautuneet koko kirjaan hyvinkin tasaisesti.

Diagrammi 17 – määrä 2/2



Tämän diagrammin sanoista yleisin on PALJON. Se ei esiinny aineistossa kovinkaan tasaisesti; sillä on kolme suurta piikkiä ja pari pienempää. Useassa kappaleessa sen määrä on kuitenkin vain 0 tai 1. KAIKKI sen sijaan on miltei koko kirjan ajan tasainen, joskin sillä on suuret piikit kappaleissa 17 ja 22. VÄHÄN esiintyy kirjassa hyvinkin tasaisesti.

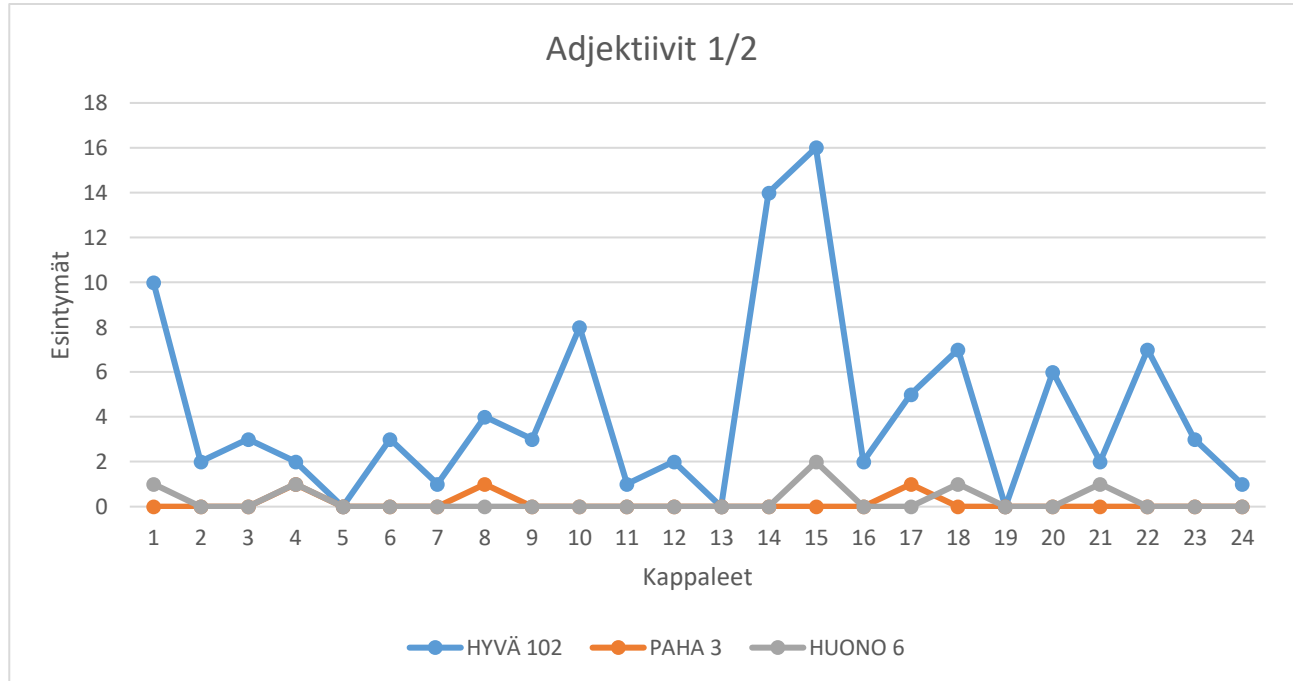
Kaikkien kolmen sanan jakaumaluku on varsin suuri: KAIKKI esiintyy 18 kertaa, PALJON 21 ja VÄHÄN 19. Ne ovat siis nähdäkseni hyvin edustettuina kirjassa.

Taulukko 4 – määrä

	Määrä	Jakauma	Määrä/jakauma
YKSI	34	13	2,62
KAKSI	32	14	2,21
MUUTAMA	23	14	1,64
JOKU	1	1	1
KAIKKI	46	18	2,56
PALJON	65	21	3,10
VÄHÄN	36	19	1,89

6.4.4 Adjektiivit

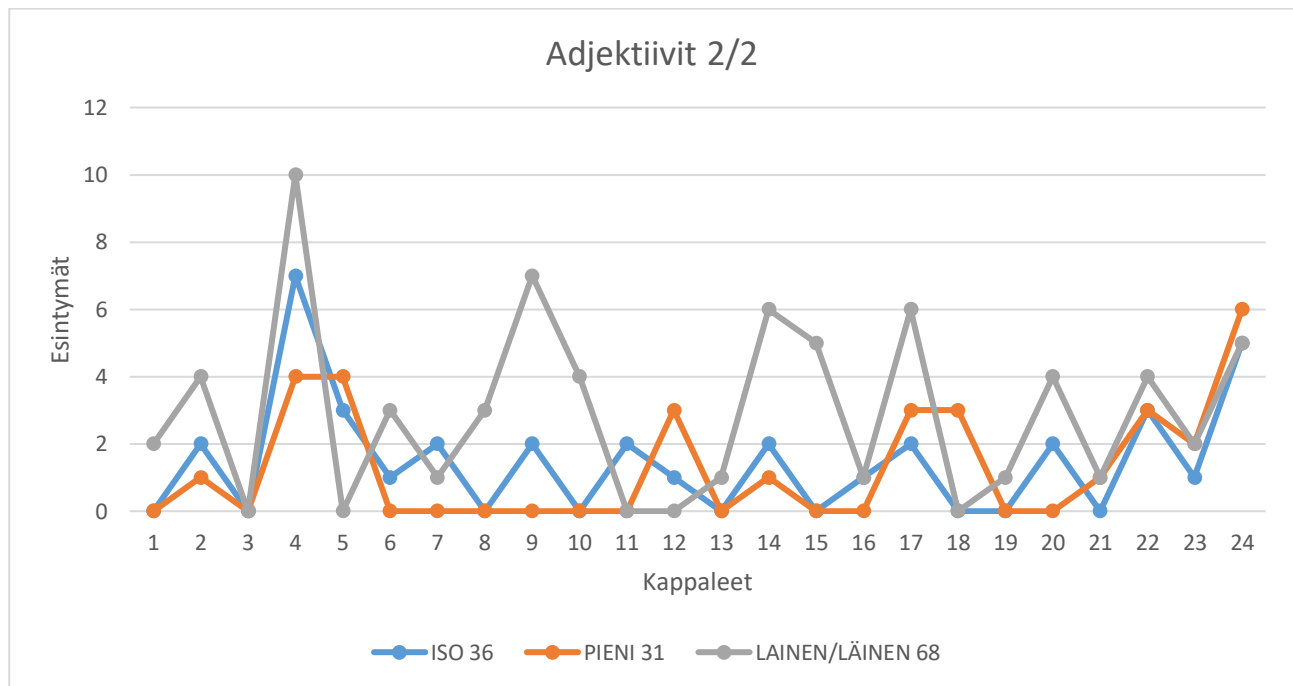
Diagrammi 18 – adjektiivit 1/2



Ensinnäkin diagrammista näkyy se, miten paljon enemmän kirja sisältää alkusanaa HYVÄ kuin sanoja PAHA tai HUONO. Kahden jälkimmäisen osalta ei niiden pienen määrän vuoksi voi puhua myöskään minkäänlaisista piikeistä. Sen sijaan alkusanalla HYVÄ on kolme kappaletta, joissa se on erityisen vahvasti esillä: 1, 14 ja 15. Kappale yksi on erityisen poikkeava sen ympäristön perusteella: Alkusanan HYVÄ vahva edustus selittynee esimerkiksi sillä, että kappaleessa opetellaan mm. kysymään ”Mitä kuuluu?”. Tähän luonnollinen vastaus on tietenkin ”hyvää”. Kappaleissa 14 ja 15 ei ole mitään yksittäistä selkeää syytä, miksi HYVÄ on niissä niin yleinen.

Alkusanojen PAHA ja HUONO jakaumasta ei ole paljoa sanottavaa. Ne ovat vähäisestä määrästäan huolimatta kirjassa tasaisesti, eli ne eivät ole pakkautuneet yhteen kappaleeseen tai asiayhteyteen. HYVÄ sen sijaan on läsnä 21:ssä kirjan kappaleessa, mikä tekee siitä sekä runsaan että jokseenkin tasaisesti esiintyvän alkusanan.

Diagrammi 19 – adjektiivit 2/2



Tämän diagrammin sanoista päätte -LAINEN/-LÄINEN on selkeästi yleisin, mutta ISO ja PIENI ovat myös melko yleisiä. Diagrammista näkyy, että -LAINEN/-LÄINEN esiintyy kirjassa jokseenkin tasaisesti, mutta saa muutama isompaa piikkiä kappaleissa 4, 9, 14, 17 ja 24. Alkusana ISO puolestaan esiintyy kirjassa jokseenkin tasaisesti alkusanaan PIENI verraten. Alkusanan ISO käyrä koostuu useista pienistä piikeistä, mutta PIENI saa usein arvon 0 ja muuten sen esiintymät koostuvat muutamasta kohtalaisen suuresta piikistä. Kappaleen neljä kohdalla oleva kaikkien käyrien korostuma selittyy sillä, että kyseisessä kappaleessa opetetaan juurikin adjektiiveja.

Sanojen jakaumaluvuista on myös löydettävissä samankaltaisia havaintoja: -LAINEN/-LÄINEN on läsnä 20:ssä kappaleessa, ISO 15:ssä ja PIENI vain 11:ssä.

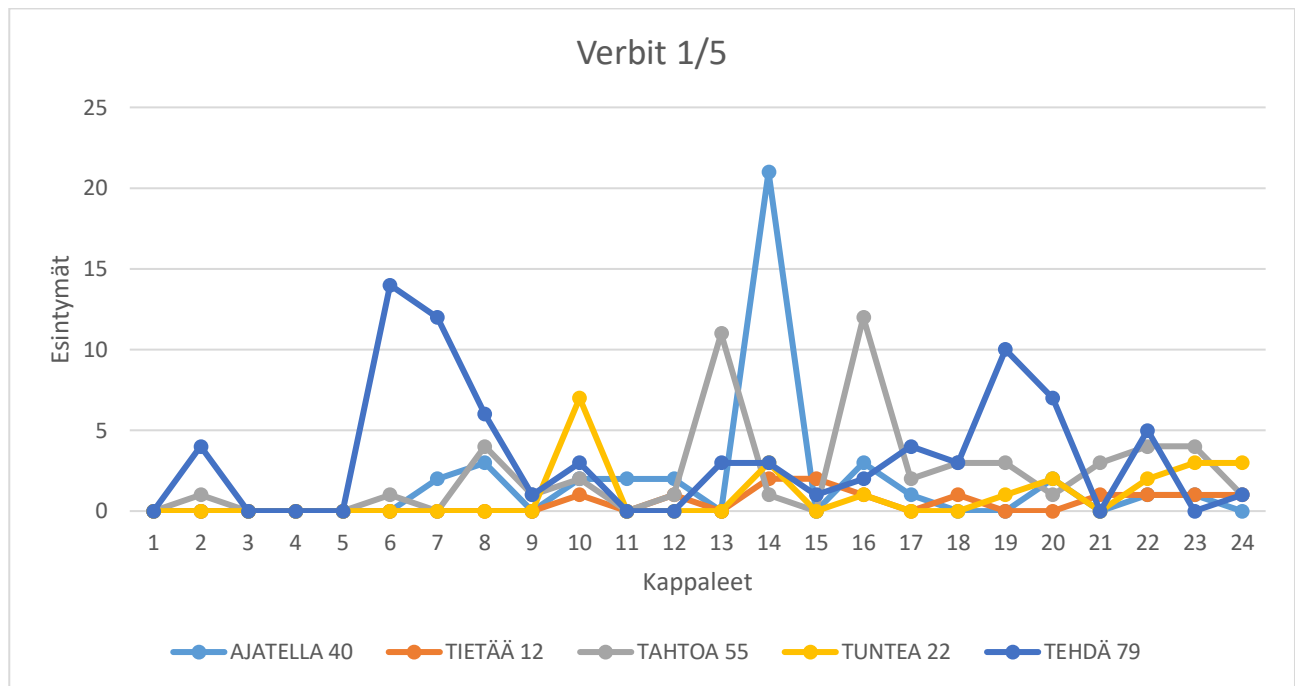
Taulukko 5 – adjektiivit

	Määrä	Jakauma	Määrä/jakauma
HYVÄ	102	21	4,86
PAHA	3	3	1
HUONO	6	5	1,2
ISO	36	15	2,4

PIENI	31	11	2,81
LAINEN/LÄINEN	68	20	3,4

6.4.5 Verbit

Diagrammi 20 – verbit 1/5

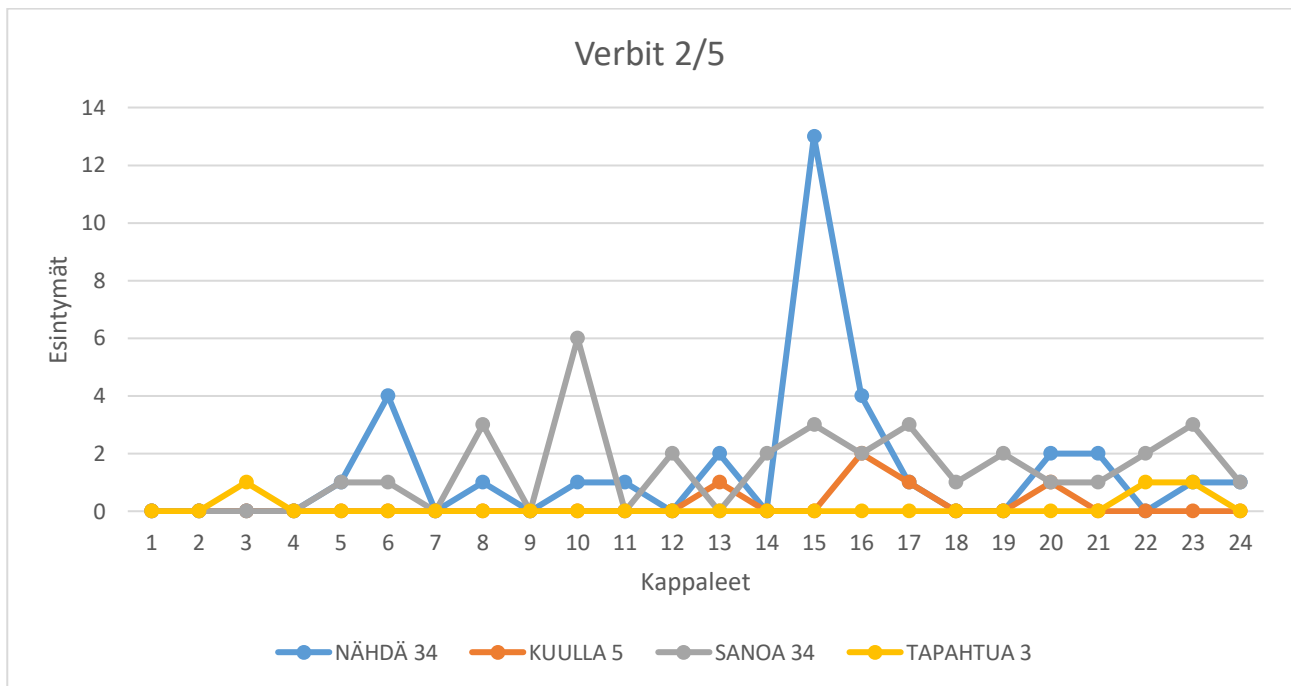


Tästä diagrammista on nähtävissä, että alkusanat TEHDÄ ja AJATELLA saavat aineistossani pari suurta piikkiä. Alkusanalla TUNEA puolestaan on yksi pienehkö piikki ja alkusanalla TAHTOA kaksi. TIETÄÄ esiintyy aineistossa harvaksen.

Alkusanan TEHDÄ kappaleen 6 piikin selittää se, että kyseisessä kappaleessa opetellaan lausetta ”Mitä sinä teet?”. Kappaleessa 14 puolestaan opetellaan kertomaan mielipide, mikä osaltaan selittää alkusanan AJATELLA piikkiä. Sanan TAHTOA kappaleen 13 piikkiä selittää osaltaan tehtävä, jossa opetellaan tahtomista kuvaavia lauseita ja kappaleen 16 piikkiä sanastollinen osa-alue, jossa opetellaan kohteliasta pyyntöä tai kysymystä: ”Haluaisin kysyä/pyytää.”

Sanoista TAHTOA esiintyy aineistossa jokseenkin taajaan, 17:ssä kappaleessa, TEHDÄ 16:ssä. AJATELLA puolestaan on läsnä 11:ssä kappaleessa, TIETÄÄ 10:ssä ja TUNTEA 8:ssä.

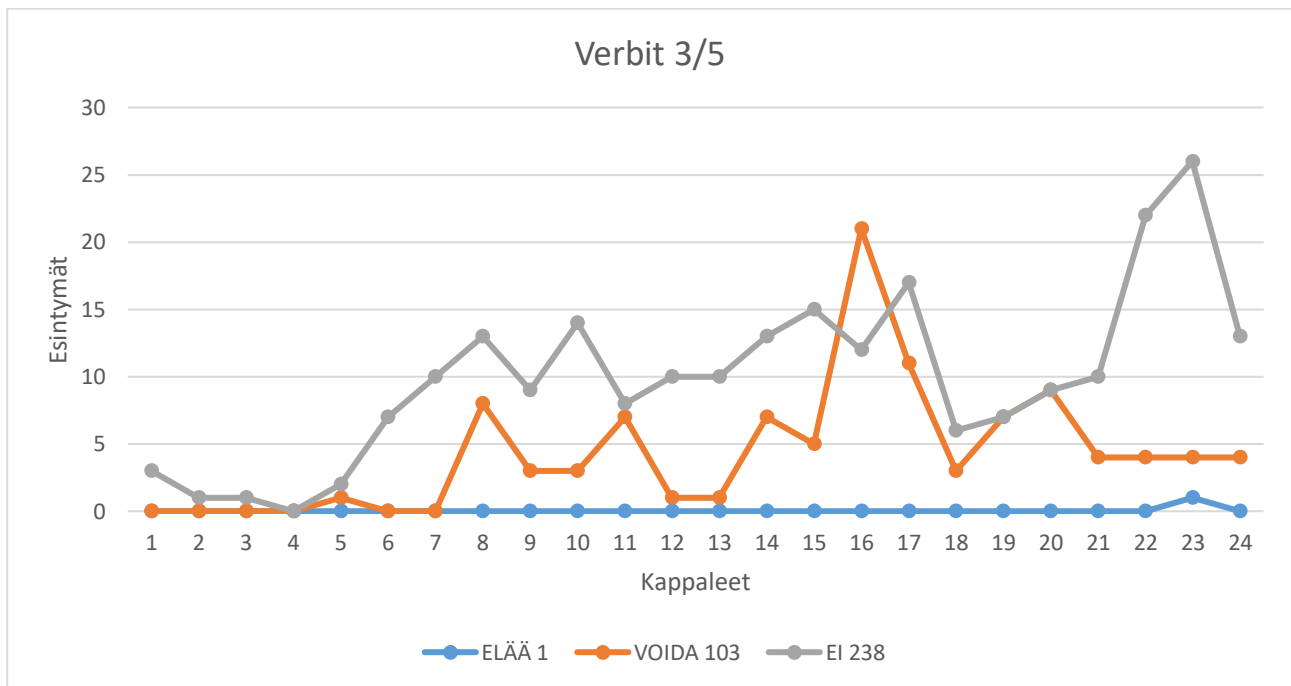
Diagrammi 21 – verbit 2/5



Tästä diagrammista on nähtävissä yksi suuri piikki: Sana NÄHDÄ esiintyy kappaleessa 15 13 kertaa, eli yli kolmasosan sen kaikkien esiintymistä. Tämä selittyy sillä, että kyseisessä kappaleessa opetellaan tapaamisen sopimista: ”Missä ja milloin nähdään?” Tämän lisäksi SANOA saa pienen piikin kappaleessa 6, mutta se ei mielestäni ole merkittävä.

Noin muuten sanat esiintyvät aineistossa nähdäkseni jokseenkin tasaisesti: NÄHDÄ esiintyy 13:ssa kappaleessa ja SANOA 16:ssa. KUULLA puolestaan esiintyy 4:ssä ja TAPAHTUA 3:ssa kappaleessa.

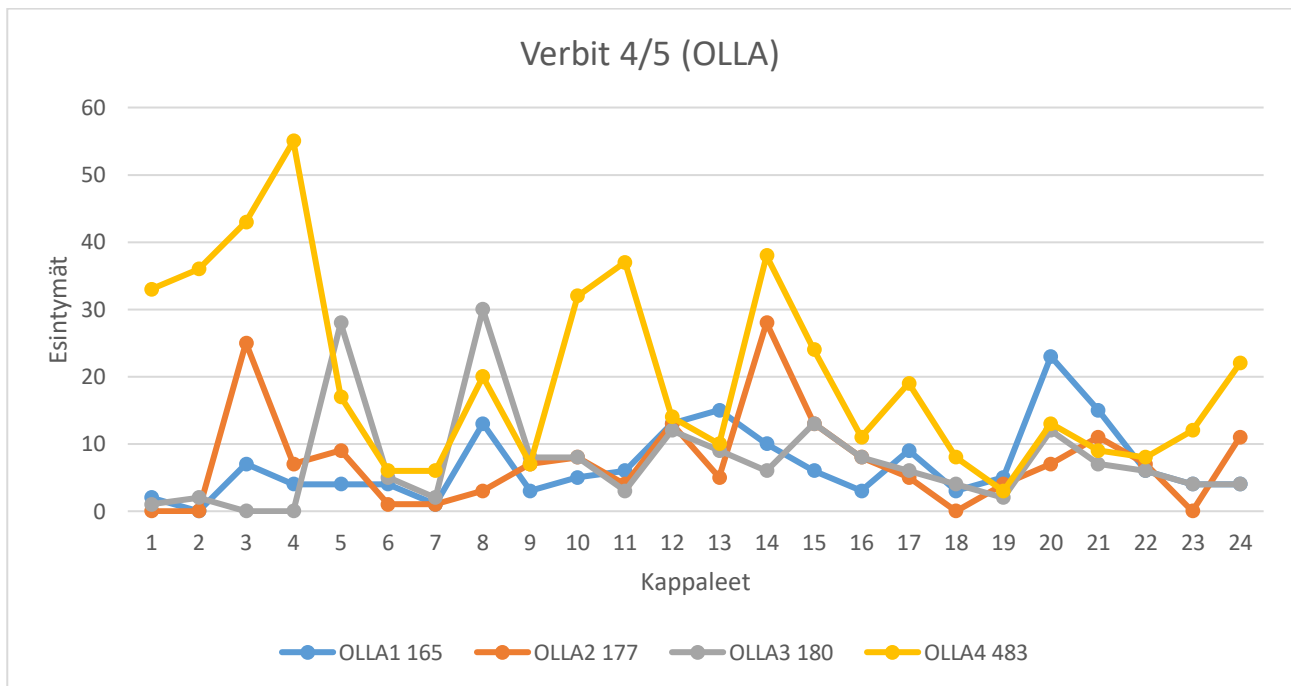
Diagrammi 22 – verbit 3/5



Tässä diagrammissa voidaan huomata se, että vaikka VOIDA ja EI ovat molemmat erittäin yleisiä aineistossani, molemmat tuodaan syötteeseen kunnolla vasta kappaleessa 6. Tästä eteenpäin molemmat ovat aineistossa edustettuina laajasti ja melko tasaisesti. VOIDA tosin saa yhden suuren piikin kappaleessa 16 ja EI kappaleissa 22 ja 23. Sanan VOIDA piikki voidaan osittain selittää kappaleen 16 teemalla *kohteliaat pyynnot*. Alkusanan EI kappaleen 22 piikkiin ei ole mitään selvää syytä, mutta kappaleessa 23 muutetaan negatiivisia lauseita positiivisiksi, mikä selittää sanan EI yleisyyttä.

Alkusana VOIDA esiintyy kirjassa melko taajaan, 18:ssä kappaleessa. EI puolestaan on erittäin taajaan läsnä, se esiintyy 23:ssä kappaleessa.

Diagrammi 23 – verbit 4/5



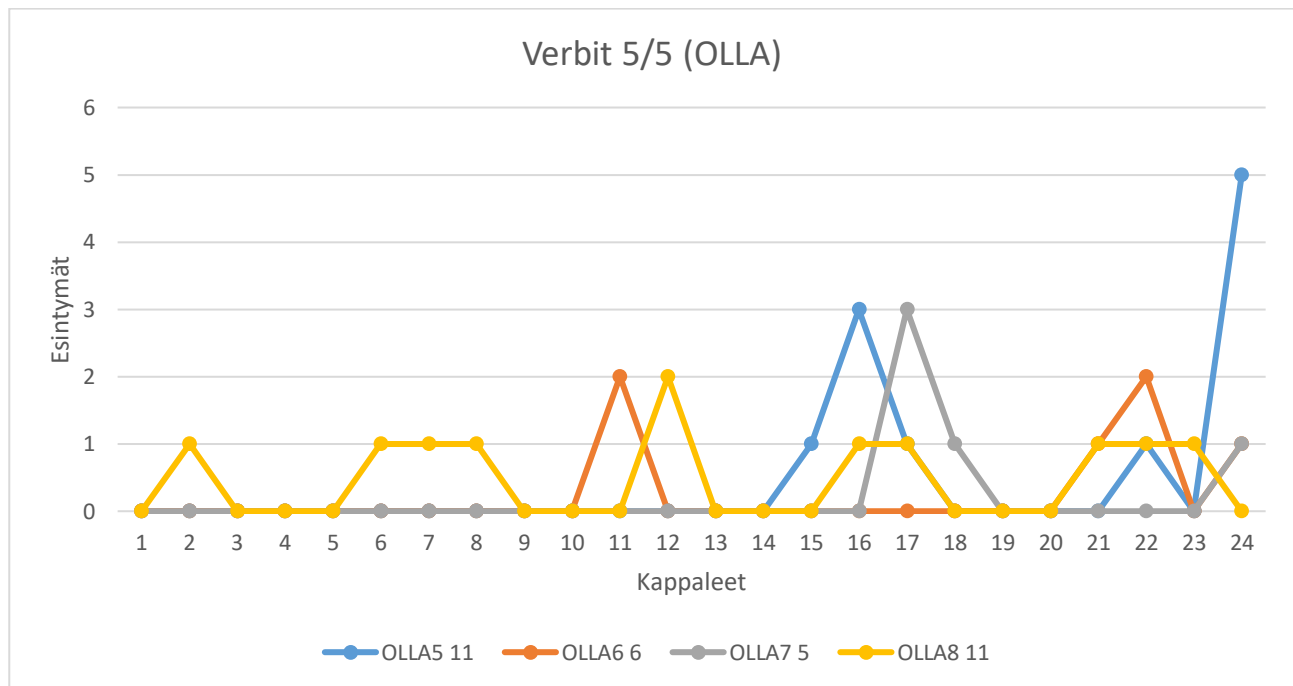
Tämä diagrammi sisältää ne *olla*-verbin neljä funktiota, jotka sisältyvät myös Vanhatalon, Tissarin ja Idströmin (2014) sekä Vanhatalon ja Tissarin (2017) ehdotuksiin suomenkieliseksi NSM-käännökseksi. Ne ovat kaikki yleisimpien alkusanojen joukossa, ja OLLA4 on selkeästi yleisin yksittäinen alkusana. OLLA4-alkusanan yleisyys näkyy mm. siinä, että se on toinen kahdesta alkusanasta, joiden jakaumaluku on 24, eli jotka ovat läsnä kaikissa kirjan kappaleissa. Toinen on HÄN. Lisäksi vain viidessä kappaleessa OLLA4:n määrä on alle 10.

Vaikka OLLA4 esiintyy aineistossa jokseenkin epätasaisesti, yksittäisiä syitä näihin piikkeihin on turha etsiä. Tämä selittyy mm. sillä, että *olla*-verbi on suomen kielen yleisin sana (Saukkonen ym. 1979) sekä sillä, että OLLA4 yhdistää kolme eri funktiota, jotka olen selvittänyt luvussa 5.

Olla-verbin yleisyyteen liittyvät huomiot ovat oleellisia myös muita diagrammin alkusanoja tarkasteltaessa. Mielestäni vain alkusanan OLLA3 kaksi piikkiä kappaleissa 5 ja 8 kahdeksan ovat huomionarvoisia: Kappaleessa 5 tuodaan kirjassa oppiaineeseen omistusmuoto, ja kappale 8 soveltaa sitä paljon esimerkiksi sairauksien yhteydessä.

Myös näiden sanojen jakauma on huomattavan korkea: OLLA1 esiintyy aineistossa 23:ssa kappaleessa, OLLA2 20:ssa, OLLA3 22:ssa ja kuten sanottua, OLLA4 kaikissa. Mainittakoon vielä OLLA4-alkusanan keskimääräinen määrä niissä kappaleissa, joissa se esiintyy: 20,125.

Diagrammi 24 – verbit 5/5



Toisin kuin edellisen diagrammin, tämän diagrammin sanat, jotka ovat kaikki itse lisäämiäni, ovat hyvinkin marginaalisia. OLLA5 kuvaa siis perfekt- ja pluskvamperfektimuodon yhteydessä käytettävää *olla*-verbiä, ja sen vähyys selittyy sillä, että kumpikaan näistä tempuksista ei kuulu kirjan op-
 piaineeseen. OLLA6 kuvaa nesessiivirakenteen yhteydessä olevan genetiivisubjektin kanssa esiin-
 tyvää *olla*-verbiä, joka ei sekään kuulu varsinaisesti kirjan ainekseen. OLLA7 kuvaa sellaisia *olla*-
 verbin infiniittimuotoja, jotka eivät sovi muihin *olla*-verbin funktioihin. OLLA8 puolestaan kuvaa
 sellaisia *olla*-verbin esiintymiä, joista ei voi luotettavasti sanoa, mihin kategoriaan ne kuuluvat.

Näiden sanojen jakaumaluvut ovat jokseenkin pieniä. OLLA5 esiintyy 5:ssä kappaleessa, OLLA6 4:ssä, OLLA7 3:ssa ja OLLA8 10:ssä.

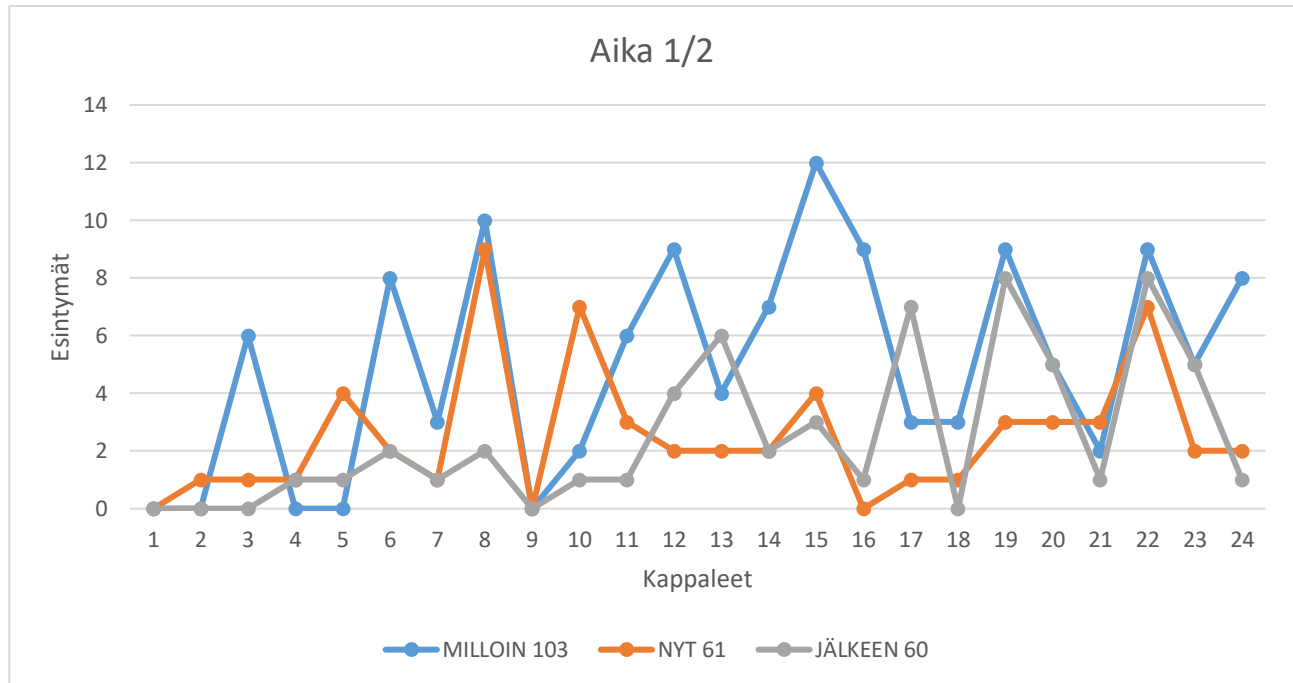
Taulukko 6 – verbit

	Määrä	Jakauma	Jakauma/määrä
AJATELLA	40	11	3,64
TIETÄÄ	12	10	1,2
TAHTOA	55	17	3,23
TUNTEA	22	8	2,75

TEHDÄ	79	16	4,94
NÄHDÄ	34	13	2,62
KUULLA	5	4	1,25
SANOA	34	16	2,125
TAPAHTUA	3	3	1
ELÄÄ	1	1	1
VOIDA	103	18	5,72
EI	238	23	10,35
OLLA1	165	23	7,17
OLLA2	177	20	8,85
OLLA3	180	22	8,18
OLLA4	483	24	20,125
OLLA5	11	5	2,2
OLLA6	6	4	1,5
OLLA7	5	3	1,667
OLLA8	11	10	1,1

6.4.6 Aika

Diagrammi 25 – aika 1/2

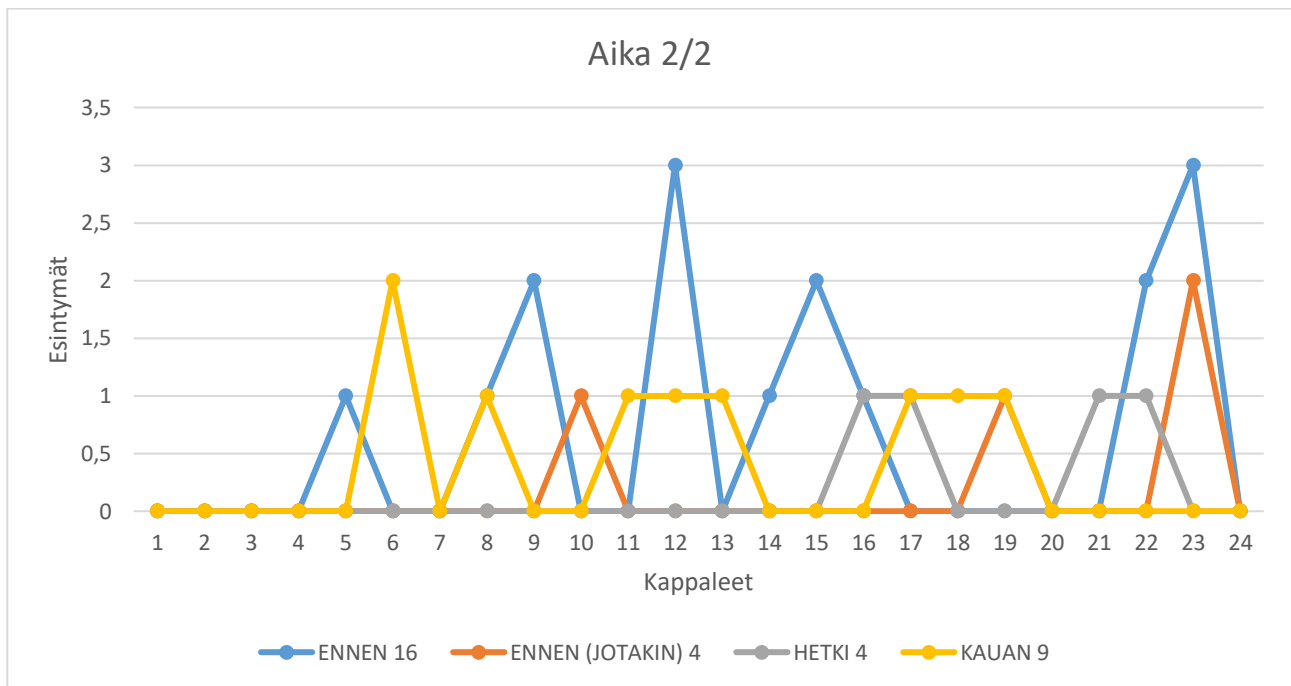


Tästä diagrammista voidaan havaita, että alkusana MILLOIN ei yleisyydestään huolimatta esiinny erityisen tasaisesti. Sen kuvaaja on jokseenkin poukkoileva, eivätkä edes yksittäiset piikit erotu. Tästä huolimatta se on laajalti läsnä aineistossa. Samat huomiot pätevät myös alkusanaan NYT.

Edellisistä poiketen alkusanan JÄLKEEN osalta voidaan huomata sen määrän nousevan kappaleesta 11 eteenpäin. Yksittäisiä piikkejä siitäkään on kuitenkin hankala löytää.

Kaikkien kolmen jakauma on varsin suuri. Milloin on läsnä 19:ssä kappaleessa, NYT 21:ssä ja JÄLKEEN 19:ssä.

Diagrammi 26 – aika 2/2



Edellisen diagrammin huomiot pätevät myös tämän diagrammin sanoihin. Vain sanat ENNEN ja KAUAN ovat tilastollisesti mainittavia, ja molemmat esiintyvät hieman poukkoillen kappaleesta 5 eteenpäin.

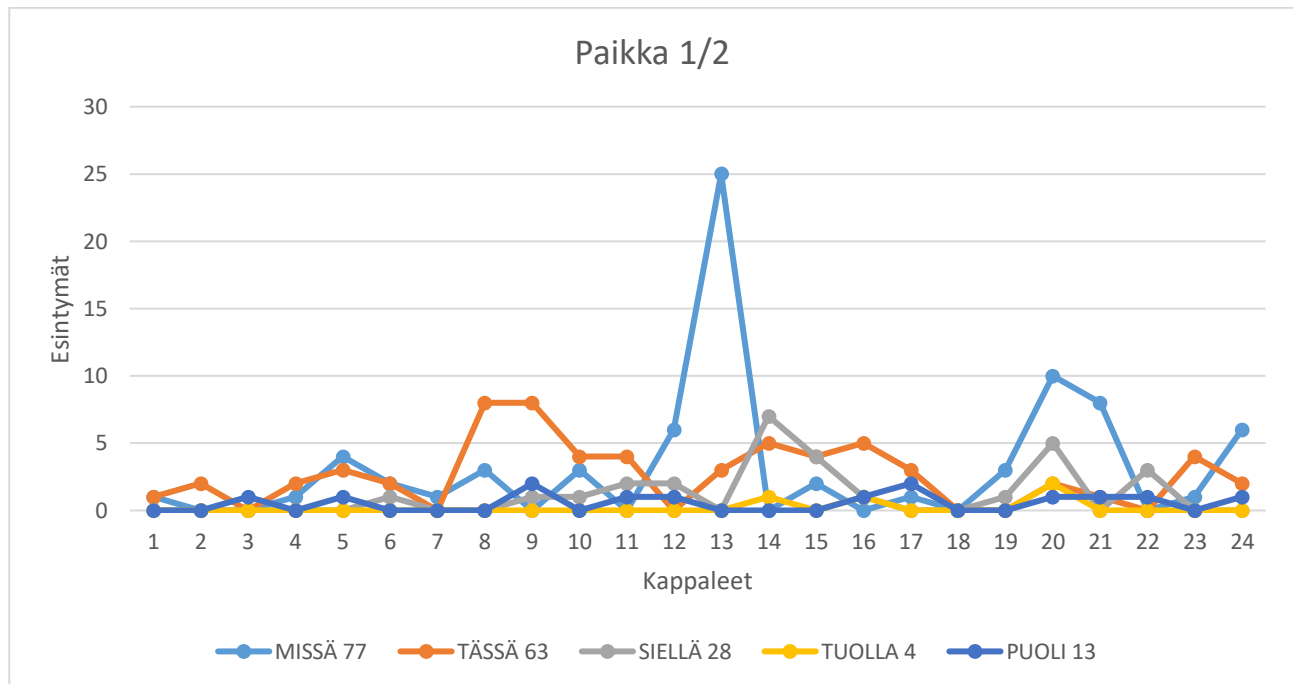
ENNEN esiintyy kirjassa 9:ssä kappaleessa, ENNEN (JOTAKIN) kolmessa, HETKI 4:ssä ja KAUAN 8:ssa.

Taulukko 7 – aika

	Määrä	Jakauma	Määrä/jakauma
MILLOIN	103	19	6,32
NYT	61	21	2,90
JÄLKEEN	60	19	3,16
ENNEN	16	9	1,78
ENNEN (JOTAKIN)	4	3	1,33
HETKI	4	4	1
KAUAN	9	8	1,125

6.4.7 Paikka

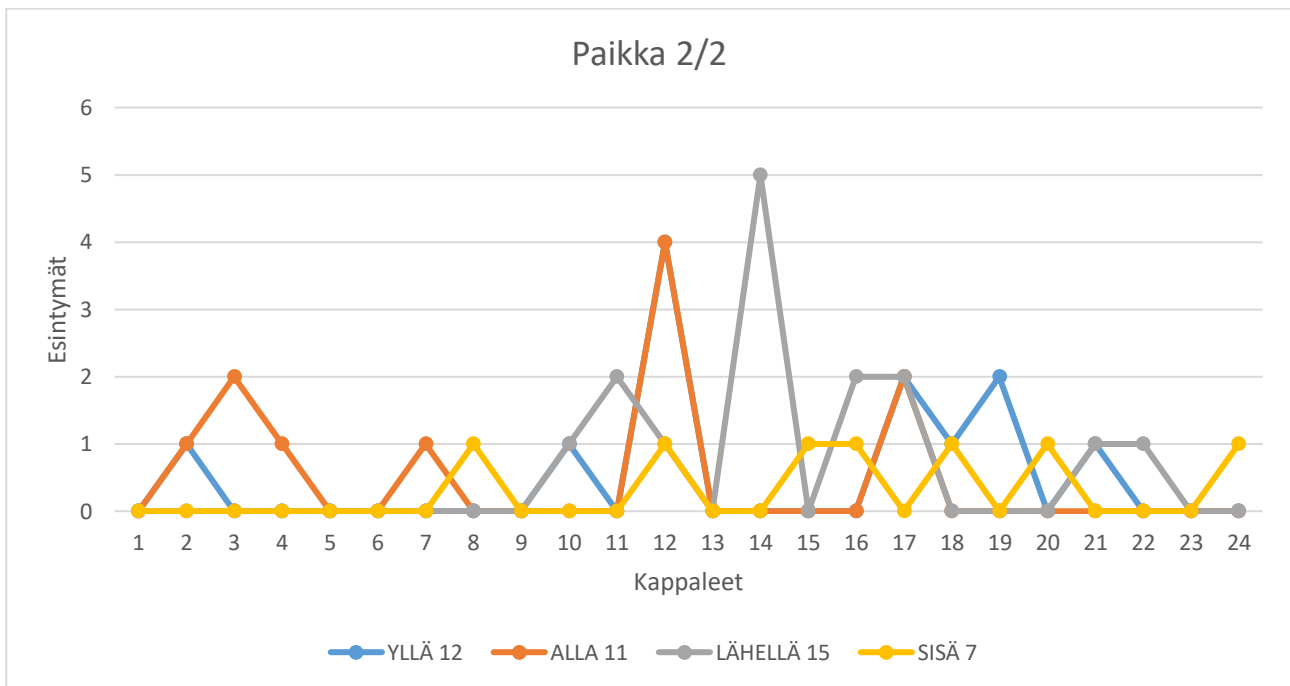
Diagrammi 27 – paikka 1/2



Selkein huomioitava asia tässä diagrammissa on alkusanan MISSÄ suuri piikki kappaleessa 13. Se selittyy sillä, että kappaleessa keskitytään paljon paikan kuvaamiseen. Muuten kuvaajat ovat jokseenkin tasaiset: TÄSSÄ erottuu hieman kappaleissa 8 ja 9, ja MISSÄ saa toisen pienehkön piikin kappaleessa 20.

Sanat MISSÄ ja TÄSSÄ ovat jokseenkin taajaan edustettuina aineistossa; MISSÄ on läsnä 16:ssa kappaleessa ja TÄSSÄ 18:ssa. SIELLÄ ja TUOLLA ovat molemmat läsnä 11:ssä kappaleessa, ja TUOLLA kolmessa.

Diagrammi 28 – paikka 2/2



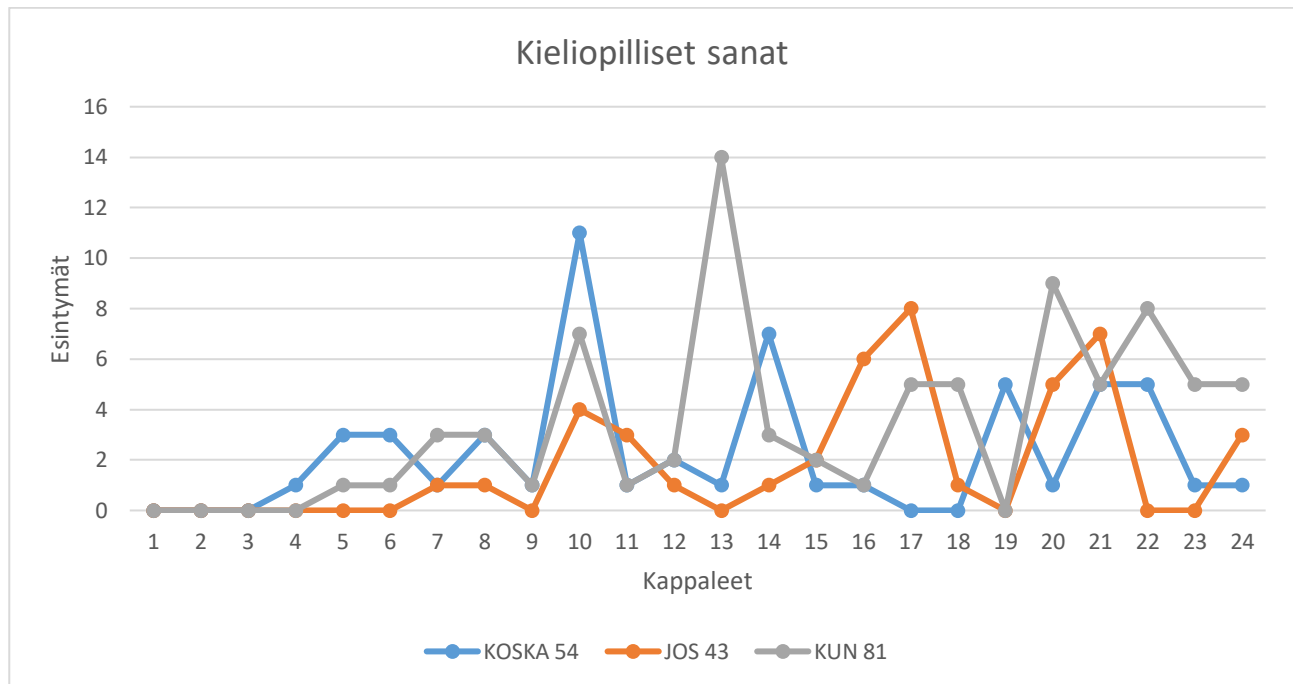
Vaikka tässä diagrammissa on suhteellisesti suuria piikkejä, eivät sanojen kokonaismäärät nouse sellaisiksi, että näillä piikeillä olisi tilastollista merkitystä. Sanat ovat kuitenkin kohtalaisen taajaan edustettuina, YLLÄ on läsnä 7:ssä kappaleessa, ALLA 6:ssa, LÄHELLÄ 8:ssa ja SISÄ 7:ssä.

Taulukko 8 – paikka

	Määrä	Jakauma	Määrä/jakauma
MISSÄ	77	16	4,83
TÄSSÄ	63	18	3,5
SIELLÄ	28	11	2,55
TUOLLA	4	3	1,33
PUOLI	13	11	1,18
YLLÄ	12	7	1,71
ALLA	11	6	1,83
LÄHELLÄ	15	8	1,86
SISÄ	7	7	1

6.4.8 Kieliopilliset sanat

Diagrammi 29 – kieliopilliset sanat



Tässä diagrammissa huomionarvoista ovat alkusanan KUN kohtalaisen suuri piikki kappaleessa 13 ja alkusanan KOSKA kappaleessa 10. Sanan KUN piikki selittyy osittain kappaleessa 13 olevilla tehtävillä, joiden tarkoitus on vastata kysymyksiin tyyppiä ”Mitä sinä teet, kun...?”. Kappaleessa 10 puolestaan käsitellään jonkin verran tuntemuksien ilmaisuja lauseilla, jotka ovat tyyppiä ”Minua harmittaa, koska...”

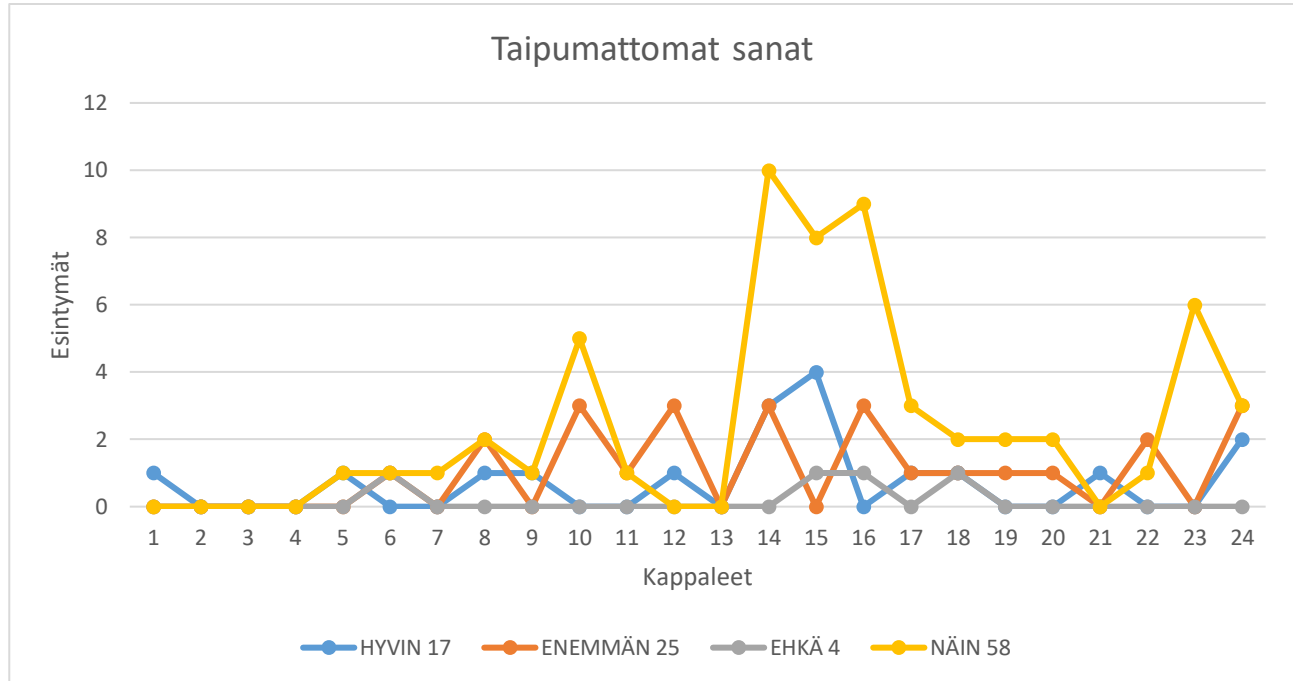
KOSKA ja KUN ovat jokseenkin taajaan läsnä aineistossa. Molemmat esiintyvät 19:ssä kappaleessa. Alkusana JOS puolestaan esiintyy 13:ssa kappaleessa.

Taulukko 9 – kieliopilliset sanat

	Määrä	Jakauma	Määrä/jakauma
KOSKA	54	19	2,84
JOS	43	13	3,31
KUN	81	19	4,26

6.4.9 Taipumattomat sanat

Diagrammi 30 – taipumattomat sanat



Tämän diagrammin kuvaajat ovat jokseenkin tasaiset lukuun ottamatta sanan NÄIN esiintymiä kappaleissa 14, 15 ja 16. Tälle ei kuitenkaan kappaleissa ole mitään ilmeistä syytä. Mielestäni on kuitenkin mielenkiintoista, että sana EHKÄ esiintyy aineistossa vain neljä kertaa.

NÄIN esiintyy aineistossa taajaan. Se on läsnä 17:ssä kappaleessa. HYVIN puolestaan on 11:ssä ja ENEMMÄN 13:ssa. Jokainen sanan EHKÄ esiintymä puolestaan on eri kappaleessa, eli sen jakauma on 4.

Taulukko 10 – taipumattomat sanat

	Määrä	Jakauma	Määrä/jakauma
HYVIN	17	11	1,55
ENEMMÄN	25	13	1,92
EHKÄ	4	4	1
NÄIN	58	17	3,41

6.4.10 Kappaleiden top3-listat

Tähdentääkseni tarkastelua esitän kunkin kappaleen osalta alkusanojen top3-listan. Tämän avulla voidaan siis tarkastella sitä, mikä alkusana on yleisin missäkin kappaleessa, mikä on hyödyllinen tieto siksi, että sen perusteella voidaan havainnoida sitä, onko jokin sana erityisen yleinen kaikissa kappaleissa, vai onko tässä vaihtelua. Vertaamalla näin saatua tietoa listaan koko kirjan yleisimmistä alkusanoista voidaan myös kartoittaa kyseisten sanojen varianssia: Karrikoituna esimerkkinä tilanteesta, jossa asetelma häiriöityisi erikoisten kerrostumien myötä, voidaan tarjota vaikka hypoteettinen kappale, jonka tehtävissä alkusana MINÄ esiintyy 200 kertaa. Tämä nostaisi alkusanan MINÄ kirjan yleisimpien joukkoon, mutta se ei kertoisi siitä mitään, kuinka laajalla rintamalla kyseinen alkusana tuodaan mukaan oppijoiden syötteeseen. Tämä top3-asetelma verrattuna alkusanojen määriin koko kirjassa siis lähestyy jo yksityiskohtaisempaa, yksittäisiin alkusanoihin keskittyvää tarkastelua, johon kiinnitän enemmän huomiota seuraavassa luvussa.

Tarkentaakseni edellisten alalukujen tarkastelua esitän vielä jokaisen kappaleen top3-listan siitä, mikä on kyseisen kappaleen yleisin alkusana. Tämä auttaa näkemään kokonais kuvaa siitä, miten mikäkin alkusana tosiaan jakautuu kirjaan: Se, että jonkin alkusanan jakautumista kuvaavassa kuvaajassa on piikki, kertoo vain sen suhteellisen suuresta määrästä kyseissä kohdassa suhteessa sen muihin esiintymiin, ei niinkään kokonaiskuvan kannalta.

Vertaamalla top3-listauksen tietoja kyseisten alkusanojen kokonaismäärään ja jakaumalukuun aineistossa pystyn rakentamaan suhteellisen kattavan kuvan siitä, miten nämä sanat kirjassa esiintyvät. Taulukko jokaisen kappaleen kolmesta yleisimmästä alkusanasta on seuraavanlainen:

Taulukko 11 – alkusanojen top3 kappaleittain

	1.	2.	3.
Kappale 1	MINÄ	OLLA4	LAINEN/LÄINEN
Kappale 2	OLLA4	SE	PALJON
Kappale 3	OLLA4	OLLA2	OLLA1
Kappale 4	OLLA4	SINÄ	LAINEN/LÄINEN
Kappale 5	OLLA3	OLLA4	HÄN
Kappale 6	MINÄ	SINÄ	HÄN

Kappale 7	HÄN	TEHDÄ	EI
Kappale 8	OLLA3	OLLA4	MINÄ
Kappale 9	MINÄ/EI	TÄSSÄ/OLLA3	
Kappale 10	OLLA4	MINÄ	SINÄ
Kappale 11	OLLA4	MINÄ	SINÄ
Kappale 12	HÄN	OLLA4	OLLA1
Kappale 13	MISSÄ	SINÄ	OLLA1
Kappale 14	OLLA4	OLLA2	AJATELLA
Kappale 15	OLLA4	TÄMÄ	HYVÄ
Kappale 16	VOIDA	TÄMÄ	EI
Kappale 17	HÄN	TÄMÄ	OLLA4
Kappale 18	HÄN	TÄMÄ	OLLA4
Kappale 19	TEHDÄ	MILLOIN	JÄLKEEN
Kappale 20	OLLA1	SINÄ	OLLA4
Kappale 21	MINÄ	HÄN	OLLA1
Kappale 22	EI	MINÄ	TÄMÄ
Kappale 23	EI	MINÄ	SINÄ
Kappale 24	OLLA4	EI	OLLA2

Edellisen taulukon sanojen osuus kappaleiden top3-listassa on siis seuraava:

Taulukko 12 – alkusanojen top3 sanoittain

	Top3-si- joitukset	1. sijat	2. sijat	3. sijat	kokonais- määrä	jakauma- luku
OLLA4	15	8	4	3	483	24
MINÄ	9	4	4	1	282	22
HÄN	7	4	1	2	223	24
SINÄ	7	0	4	3	225	23
OLLA1	5	1	0	4	165	23
TÄMÄ	5	0	4	1	143	21
EI	4	3	1	0	238	23
OLLA3	3	2	1	0	180	22

OLLA2	3	0	2	1	177	20
TEHDÄ	2	1	1	0	79	16
LAINEN	2	0	0	2	68	20
MISSÄ	1	1	0	0	77	16
VOIDA	1	1	0	0	103	18
SE	1	0	1	0	119	23
MILLOIN	1	0	1	0	120	19
TÄSSÄ	1	0	1	0	63	18
AJATELLA	1	0	0	1	40	11
HYVÄ	1	0	0	1	102	21
JÄLKEEN	1	0	0	1	60	19
PALJON	1	0	0	1	65	21

Tästä taulukosta voi huomata, että jokainen alkusana, jolla on yli 100 esiintymää kirjassa, on myös vähintään yhden kappaleen top3-listassa. OLLA4 on vähemmän yllättävästi listan kärjessä, ja perässä ovat persoonapronominit MINÄ ja SINÄ sekä alkusana HÄN, jonka alaisuuteen kuuluvat myös sanat IHMINEN, HENKILÖ ja JOKU.

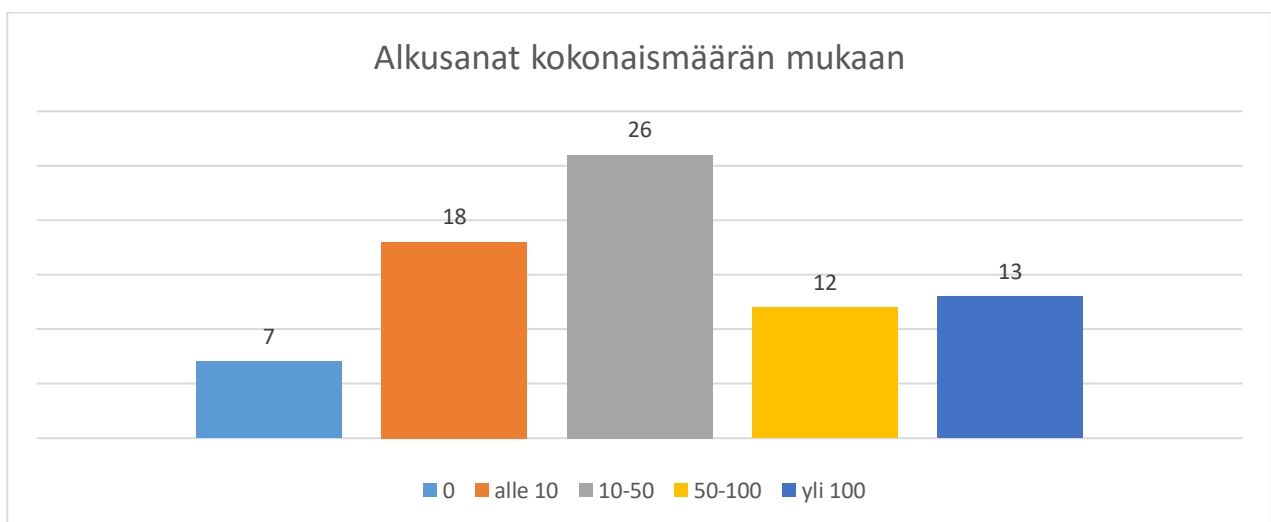
Ainoa alkusana, jolla on alle 50 esiintymää kirjassa, mutta joka on tässä listauksessa, on AJATELLA. Se ylipäättään esiintyy vain 11:ssä kappaleessa. Se, että se kuuluu tähän listaukseen, tarkoittaa, että sanan AJATELLA piikki kappaleessa 14 on jopa tilastollisesti merkittävä. Tämä tarkoittaa, että yksittäinen kimppu esiintymiä voi nostaa yhden alkusanan tilastollisesti merkittäväksi kokonaisen kirjan tasolla. Samaa huomiota tukee se, että MISSÄ on kappaleen 13 yleisin alkusana, vaikka muuten se ei pääse edes top3-listaukseen: Kappaleen 13 piikki sanan MISSÄ esiintymissä tekee siitä koko kirjan tasolla merkittävän alkusanan.

6.5 Oleellisimmat tulokset

Edeltävien lukujen perusteella voin siis todeta joitain asioita. Ensinnäkin moni alkusanoista on erittäin taajaan ja viljalti edustettuna kirjassa. Ne ovat sekä runsaita, että yleisiä pitkin koko kirjan. Tällaisia sanoja ovat esimerkiksi OLLA4, jolla on miltei 500 esiintymää kirjassa, ja joka esiintyy kirjan jokaisessa kappaleessa ja MINÄ, jolla esiintymiä on 282, ja joka esiintyy kaikissa kirjan kappaleissa paitsi kahdessa.

Kokonaisuudessaan 25 alkusanaa esiintyy kirjassa yli 50 kertaa. Toisaalta, 7 alkusanaa ei esiinny koko kirjassa lainkaan ja 18 alkusanaa saavat esiintymiä vain alle 10. Kun tämän informaation esittää taulukkona nähdään, että se noudattaa karkeasti Gaussin käyrää:

Diagrammi 31 – alkusanat määrien mukaan



Kun tarkastellaan kirjaa sen kolmen osa-alueen, tekstien, tehtävien ja sanastollisten kokonaisuuksien mukaan, voidaan ensin todeta, että kokonaismäärältään alkusanoja on enemmän teksteissä ja tehtävissä kuin sanasto-osuuksissa. Suhteutettuna kyseisten osa-alueiden kokonaissanamäärään asia ei kuitenkaan ole näin yksiselitteinen: Tekstit ja tehtävät ovat myös suhteelliselta alkusanamäärältään jokseenkin tasoissa tekstien sisältäessä alkusanoja hieman enemmän, mutta sanastollisten kokonaisuudet vaihtelevat suhteellisen alkusanamääränsä mukaan huomattavasti: Välillä alkusanojen prosentuaalinen määrä on suurempi kuin teksteissä ja tehtävissä, välillä se on nolla. Kokonaisuutena alkusanoja kuitenkin on eniten sanastoa kuvaavilla osa-alueilla sanamääriin suhteutettuna.

Tämä kuvaa näitä sanastoa opettavia osa-alueita: Ne sisältävät sanoja määrällisesti vähän, jolloin alkusanojen ollessa vahvasti edustettuina koko osa-alue nostaa profiiliaan tilastoinnissa. Toisaalta sanat esiintyvät näissä osa-alueissa kimppuina, jolloin voidaan havaita piikki myös alkusanojen suhteellisessa määrässä.

Sama asia on huomattavissa myös eri osa-alueiden varianssia kuvaavasta kuvaajasta: Alkusanoja esiintyy teksteissä ja tehtävissä jokseenkin saman verran, mutta sanasto-osa-alueissa niiden varianssi on huomattavasti pienempi. Toisin sanoen eri sanojen määrä on teksteissä ja tehtävissä jokseenkin sama, mutta sanasto-osa-alueissa on eri sanojan huomattavasti vähemmän.

Kirjan eri kappaleita tarkasteltaessa alkusanoja on jonkin verran enemmän kirjan loppupuoliskolla kuin sen alussa. Tämä kuitenkin selittyy hyvin pitkälti sillä, että kirjan loppupuolella myös kappaleet alkavat olla pitempiä: Niillä on enemmän tekstiä ja enemmän tehtäviä. On tosin syytä mainita, että aivan kirjan lopussa sanastollisia osa-alueita ei ole lainkaan.

Alkusanojen kokonaismäärän selittymisen sillä, että tekstien pituus kasvaa kirjan loppua kohden, voimme päätellä siitä, että alkusanojen suhteellinen määrä pysyy yllättävänkin tasaisena koko kirjan ajan. Keskimäärin tekstit, tehtävät ja sanastolliset osa-alueet eivät siis muutu sen perusteella, kuinka paljon alkusanoja niissä suhteellisesti on.

Selkeä muutos kirjassa on kuitenkin nähtävissä varianssissa, eli siinä, kuinka moni alkusana on läsnä missäkin kappaleessa. Kirjan alussa varianssi on pieni, mutta se kasvaa huomattavasti kirjan loppua kohden. Tämä tarkoittaa sitä, että kirjan sisältö käy monipuolisemmaksi myös alkusanojen suhteen kirjan edetessä. Alkusanojen määrä ei siis niinkään lisäännä, mutta se, kuinka montaa alkusanaa aineksessa käytetään, lisääntyy.

Yksityiskohtaisesti tarkasteltuna eri alkusanojen esiintymissä on suuriakin eroja. Jotkut ovat läsnä kirjassa tasaisesti, toiset puolestaan esiintyvät tilastoissa jopa yhden ainoan kertymän takia. Katson, että alkusanojen runsas, mutta tasainen esiintyminen tarjoaa paremmat lähtökohdat niiden oppimiselle kuin se, että ne esiintyvät yksittäisessä kohdassa erityisen tiheään. Toisaalta, jos kiinnitetään selkeästi huomiota tiettyyn alkusanaan, sen esiintyminen tiheästi jossakin kohtaa ei välttämättä ole huono asia: Tällöin voidaan kiinnittää huomiota sanan erilaisiin käyttökohteisiin ja niihin tilanteisiin, joissa se voi esiintyä. Yhtä oikeaa tapaa ei siis ole, ja perusteluita löytyy varmasti molemmille.

7. PÄÄTELMÄT JA POHDINNAT

Tässä luvussa esittelen vielä tutkimuskysymykseni ja vastaukset, jotka olen niihin kyennyt antamaan. Kuvailen myös, mitä tulokset kertovat tutkimusongelmasta ja alkusanoista ylipäätään. Tämän jälkeen esittelen ajatuksiani siitä, miten nykyinen tilanne vastaa oppimisen kannalta ideaalia tilannetta, ja mitä sille olisi mahdollisesti tehtävissä. Lopuksi pohdin vielä, miten tutkimusta aiheen tiimoilta voisi jatkaa.

Kuten luvussa 1 kirjoitin, tutkimuskysymykseni olivat:

1. Mitkä NSM-alkusanat esiintyvät valitussa S2-oppikirjassa ja kuinka yleisiä ne ovat?
2. Millaisissa yhteyksissä esiintyvät sanat ovat ja miten niitä käytetään? Miksi?

Lisäksi esittelin johdantoluvussani kaksi kysymystä, jotka pyrin huomiomaan työn aikana:

- Mitkä alkusanat voisivat olla hyödyllisiä toisen kielen opetusta ajatellen ja mitkä eivät?
- Millaisia erityistä huomiota vaativia piirteitä suomen kielessä on NSM-alkusanoja ajatellen?

Olen mielestäni kyennyt vastaamaan ensimmäiseen kysymykseen jokseenkin kattavasti luvussa 6. Alkusanojen jakautuminen oppikirjassa on paikoin runsasta, mutta hyvin epätasaista. Kuten luvussa 6.5 kirjoitin, tasainen esiintyminen ei välttämättä ole itseisarvo, mutta jotta voitaisiin olettaa oppijan alkusanat oppivat, niiden tulisi mielestäni olla aineistossa enemmän läsnä.

Luvun 6.5 diagrammin mukaan suurin osa, 26 alkusanaa, alkusanoista saa kirjassa noin 10-50 esiintymää. Myös sellaisia alkusanoja, joita kirjassa on 50-100 tai jopa yli 100, on jokseenkin paljon. 25 alkusanaa esiintyy kirjassa ainakin 50 kertaa. Vastapainona tälle 25 alkusanaa esiintyy kirjassa vähemmän kuin 10 kertaa. Tämä on se alue, johon tulisi mielestäni kiinnittää eniten huomiota, jos aikeena on rakentaa alkusanoihin perustuvaa opetusmateriaalia.

Ne alkusanat, joita kirjassa on yli 50 ovat tämän lisäksi yleisiä suomen kielessä ylipäätään (Saukkonen ym. 1979), puhumattakaan siitä, että niiden voidaan mielestäni perustellusti Ellisin (1994) käsitettä noudattaen *unmarked*. Ei siis ole mielestäni mitään syytä epäillä, etteikö kielenopijalla olisi kaikkia avuja näiden sanojen oppimiseen. Tilanne eroaa jonkin verran niistä alkusanoista, joita kirjassa on alle 10. Vaikka nekin ovat mitä ilmeisimmin *unmarked*, viitaten Ellisin si-

teeraamaan (Bardovi-Harlig 1987) tutkimukseen, toistuva, aktiivinen syöte saattaa jopa ajaa *unmarked* piirteiden ohi oppimisen tehokkuudessa. Tästä syystä ne täytyisi ottaa huomioon myös syötettä rakennettaessa.

Vastauksena toiseen kysymykseen voin todeta, että pääasiassa ne alkusanat, jotka ylipäätään ovat kirjassa edustettuina, ovat läsnä siinä melko taajaan. Niitä siis hyödynnetään erilaisissa yhteyksissä, minkä katson palvelevan oppimista. On kuitenkin hyvä huomata, että jokunen alkusana esiintyy oikeastaan vain sille erikseen varatussa paikassa kirjassa, esimerkiksi sellaisessa kappaleessa tai osa-alueessa, joka keskittyy juuri siihen. Tälläkin on puolensa, mutta mielestäni alkusanan tällainen esiintyminen ei valjasta käyttöönään kaikkea sitä potentiaalia, joka näillä alkusanoilla voisi oppimisen kannalta olla.

Katson hypoteesieni pitäneen osittain paikkansa. Tietyt alkusanat esiintyvät kirjoissa koko ajan, toiset eivät koskaan. On myös totta, että monet alkusanat korostuvat sellaisissa kohdissa kirjassa, joissa niitä opetetaan ryppäissä. Tällaisia ovat esimerkiksi jotkut adjektiivit.

Esittämäni kysymykseen siitä, mitkä alkusanat voisivat olla hyödyllisiä oppijalle, jolle niitä haluaisin opettaa, voin esittää lähinnä vain omaa pohdintaani. Katson kuitenkin, että Saukkosen ym. Suomen kielen taajuussanasto (1979) esittäytyy tässä tärkeässä osassa: Jos jokin sana on erittäin harvinaisen kielessä yleensä, miksi se olisi erityisen hyödyllinen NSM:ään perustuvassa oppimisessä?

On kuitenkin myönnettävä, että alkusanat ovat hyvin pitkälti luonteeltaan yksinkertaisia ja todella perimmäisiltä sanoilta vaikuttavia. Tähän on kuitenkin joitakin poikkeuksia, joita olen tuonut esiin pitkin tutkielmaani: Mielestäni sanan KUOLLA ei tulisi olla alkusana, sillä nähdäkseni sen voi esittää muita alkusanoja käyttäen. Tämä tosin ei ole myöskään aukoton perustelu. Kuten Goddard (1998a) esittää, sanan määritelmä ei saisi olla monimutkaisempi kuin se itse. Jos siis haluttaisiin kirjoittaa eksplikaatio sanalle KUOLLA, tulisi pohtia sitä, mitä sanoja käyttää. Esimerkiksi *hengittää*, *sydän* ja *pysähtyä* ovat jokseenkin monimutkaisia sanoja. Toisaalta, eivät mielestäni kuitenkaan niin monimutkaisia kuin *kuolla*. Sanan kuolla voisi esittää myös negaation avulla toisesta alkusanasta ELÄÄ.

Sanaan RUUMIS pätevät mielestäni ainakin jollain tavalla samat perusteet kuin sanan KUOLLA. Lisäksi ilmaisut JONKIN AIKAA ja VÄHÄN AIKAA esiintyvät mielestäni hyvin kyseenalaisina alkusanoina.

Mitä tulee sanojen käytettävyyteen yleisesti, suurimman osan uskoisin olevan hyvin tehokkaita käytössä. Sen sijaan juurikin sana KUOLLA vaikuttaa myös tässä valossa erikoiselta. JONKIN AIKAA ja VÄHÄN AIKAA sen sijaan vaikuttavat erittäin hyödyllisiltä ilmaisuilta, vaikka niiden luonteen alkusanoina kyseenalaistankin.

Kysymys siitä, mitä asioita tulee ottaa huomioon, kun NSM:ää sovelletaan suomen kieleen on mielestäni erittäin tärkeä. Olen pyrkinyt erittelemään siihen liittyviä ajatuksiani mahdollisimman tehokkaasti erityisesti luvussa 5, mutta haluan vielä todeta, että kysymys vaatii lisätutkimusta. Suomen kieli poikkeaa monesta muusta maailman kielestä paikoin rajustikin, ja tällaisia kieltenvälisiä poikkeuksia ei missään nimessä saa laiminlyödä tutkimuksessa. Englanninkielisestä NSM:n versiosta on tehty esimerkiksi varsin kattava lista, josta näkee, millaisissa käyttöyhteyksissä mitkäkin alkusanat ovat hyväksyttäviä ja millaisissa eivät. Kyseisen listan ovat suomentaneet Vanhatalo ja Tissari (2017), ja se on työssäni liitteenä (liite 1).

Mielestäni tulokset kuvastavat alkusanojen jakautumista kieleen jonkin verran. On tosin totta, että oppikirjateksti eroaa luonnollisesta kielestä, eli tulokset eivät ole suoraan yleistettävissä. Erityisesti tehtävä- ja sanastokokonaisuudet poikkeavat luonnollisesti luonnollisesta kielestä vielä tekstikokonaisuuksia enemmän.

Sen voi kuitenkin aineistoni perusteella todeta, että jotkin alkusanat ovat huomattavasti yleisempiä kuin toiset. Niiden jakautuminen ei ole kaoottista tai sattumanvaraista, vaan uskon sen kuvastavan itse kielen rakennetta tavalla tai toisella.

Mitä tulee oppimateriaalissa vallitsevaan tilanteeseen, se on mielestäni kelvollinen. Kuten sanottua, alkusanat ovat suureksi osaksi kirjassa läsnä, mutta moni jää silti kirjasta uupumaan. Sellainen sana, joka esiintyy kirjassa vain kerran, tai edes alle kymmenen kertaa, on nähdäkseni niin heikosti edustettuna, ettei oppijan voi välttämättä olettaa sitä oppivan.

Käytännössä tämä tarkoittaa mielestäni sitä, että NSM tulisi todellakin ottaa huomioon oppikirjallisuutta ja -materiaalia suunniteltaessa. Jos oppijoille voidaan näin tarjota pääsy alkusana-kieleen, ilman edes varsinaista sanoihin keskittymistä, on se mielestäni tekemisen arvoinen kehitys. Vielä arvokkaammaksi asian tekisi se, jos lähdetäisiin kehittämään varsinaista metodologiaa, joka tähtää semanttisiin universaaleihin yhtenä opetuksen kulmakivenä.

Edellisestä huomiosta päästään tämän pohdintakappaleen viimeiseen osioon. Tämän tutkimuksen jälkeen on hyvä pohtia, miten tutkimusta ja alkusanoihin perustuvaa opetusta voisi jatkossa kehittää.

Ensimmäinen huomioni liittyy omaan aineistooni, ja etenkin sen suppeuteen. Tutkielmani on ajallisten ja tilallisten rajoitteiden vuoksi tehty vain yhdestä, joskin laajalti käytetystä, oppikirjasta. On kuitenkin huomattava, että jotta voitaisiin todella maalata yleiskuva alkusanoista S2-opetuksessa, tulisi tutkia muitakin kirjoja, internet-materiaaleja, opettajien itse tekemiä materiaaleja,

sanastoja, sanakirjoja, opetusohjelmia ja -pelejä ynnä muuta vastaavaa. Lisäksi voitaisiin ottaa yhteys opetusta järjestäviin tahoihin ja keskustella mahdollisuuksista NSM:n käytöstä opetuksessa. Näin varmistettaisiin, että NSM:ää hyödyntämään pyrkivä opetusmenetelmä olisi todellakin vakasti perusteltu.

Jos edelliset huomiot tähtäävät vallitsevan asiantilan tutkimiseen, toinen näkökulma keskittyy tulevaisuuteen: Millaisen metodin alkusanojen perusteella sitten loisi? Opetettaisiinko opiskelijoille Ellisin (1994) kuvailemia *communication strategies* -kykyjä, miten otollinen tilanne rakennettaisiin, miten testattaisiin opiskelijoiden taitoja? Mahdollisuuksia on runsaasti. Mielestäni NSM:llä on paljon sovellusmahdollisuuksia toisen kielen oppimisen saralla, enkä kiistä, etteikö oma tutkimuksenikin olisi tarjonnut joitakuuta.

Toisen kielen opetus on alati kehittyvä ala, sillä emme aina ymmärrä täysin edes kielen tai oppimisen toimintaperiaatteita. Kielen ymmärtäminen liittyy osaltaan myös ihmismielen ymmärtämiseen, joka puolestaan on vahvasti yhteydessä oppimiseen. Tästä syystä peräänkuulutan lisätutkimusta aiheen tiimoilta: Tutkimuskohde tuskin on koskaan täysin tutkittu tai aineisto koskaan kylläinen. Tämä mielestäni vain kannustaa panostamaan lisää aikaa ja työtunteja kielen ja oppimisen tutkimukseen.

8. LÄHTEET

Aikuisten maahanmuuttajien kotoutumiskoulutuksen opetussuunnitelman perusteet 2012. Opetushallitus. Juvenes Print – Tampereen yliopistopaino Oy 2012.

ARNAULD, ANTOINE – NICOLE, PIERRE 1996[1662]: *Logic, or, the Art of Thinking*. Buroker, Jill Vance (toim.). Cambrige University Press 1996.

BARDOVI-HARLIG, KATHLEEN 1987: *Markedness and salience in second language acquisition*. *Language learning* 37: 385-407.

BIRNER, BETTY: *Does the language I speak influence the way I think?* <https://www.linguisticsociety.org/content/does-language-i-speak-influence-way-i-think> (6.2.2018).

BLOOMFIELD, LEONARD 1933: *Language*. New York.

BLUMCZYNSKI, PIOTR 2013: *Turning the tide: A critique of Natuarl Semantic Metalanguage from a Language studies perspective*. *Translation studies* 2013.

BOHNEMEYER, JÜRGEN 2003: *Comments on Durst's article*. University at Buffalo. SUNY.

BROAD, CHRIS: *Speaking Japanese Fluently in Six Months. Abroad in Japan*. https://www.youtube.com/watch?v=_IOZbJ7PCPk (15.3.2018).

CHOMSKY, NOAM 1965: *Apects of the theory of syntax*. Cambrige, Mass.: MIT Press.

DE SAUSSURE, FREDINAND 1962 [1916]: *Cours de linguistique générale*. Paris. Payot. 1962.

ECKMAN, FRED 1977: *Markedness and the contrastive analysis hypothesis*. *Language learning* 27: 315-30.

ELLIS, ROD 1994: *The study of second language acquisition*. Oxford University Press. 1994.

ELLIS, ROD 1987: *Understanding second language acquisition*. Oxford University Press. 1987.

EVANS, NICHOLAS 1994. *Kayardild*. Goddard & Wierzbicka (toim.): 203-228.

FAWCETT, PETER 1997: *Translation and language. Linguistic theories explained*. St. Jerome Publishing. 1997.

GEERAERTS, DIRK 2009: *Theories of Lexical Semantics*. Oxford University Press.

GODDARD, CLIFF 1994: *Semantic and lexical universals: Theory and empirical findings*. Goddard & Wierzbicka (toim.). Amsterdam: John Benjamins. 1994.

————— 1998a: *Semantic Analysis – A Practical Introduction*, Oxford University Press.

————— 1998b: *Bad arguments against semantic primitives*. Theoretical linguistics 24/2-3. 1998.

————— 2001: Conceptual in Early Language Development – Pütz, M. & Niemeier, S. (toim.) *Applied Cognitive Linguistics I: Theory and Language Acquisition*. Berlin: Mouton de Gruyter.: 193-227.

————— 2002: The search for the shared semantic core of all languages. Goddard & Wierzbicka (toim.). *Meaning and Universal Grammar – Theory and Empirical Findings, Vol. 1*. Amsterdam: John Benjamins: 5-41.

————— 2010: *Cultural Scripts: Applications to Language Teaching and Intercultural Communication*. Studies in Pragmatics 3. Griffith University.: 105-119.

————— 2011: *Semantic analysis. A practical introduction*. Toinen, uudistettu painos. Oxford: Oxford University Press.

GODDARD CLIFF, PEETERS BERT 2006: *The Natural Semantic Metalanguage Approach: An overview with reference to the most important Romance languages*. Amsterdam, John Benjamins. 2006.

GODDARD, CLIFF – WIERZBICKA, ANNA 1994: Introducing lexical primitives. Goddard & Wierzbicka (toim.), *Semantic and Lexical Universals: Theory and Empirical Findings*. Amsterdam: John Benjamins.: 31-54.

————— 1995: *Key words, culture and cognition*. Philosophica 55, nro. 1.: 37-67.

————— 2007: Semantic primes and cultural scripts in language, F. & Palmer, G. (toim.), *Applied Cultural Linguistics*. Amsterdam: John Benjamins.: 105-124.

HONKO, MARI 2013: *Alakouluikäisten leksikaalinen tieto ja taito: toisen sukupolven suomi ja S1-verrokki*. Tohtorinväitöskirja. Tampereen yliopiston suomen kielen laitos.

HUMBOLDT, WILHELM VON 1988. *On language: The diversity of human language structure and its influence on mental development of mankind*. Käänt. Peter Heath. New York: Cambridge University Press.

HYLTENSTAM, KENNETH 1984: *The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses*. Andersen (toim.). 1984.

ITKONEN, ESA 1983: *Causality in linguistic theory*. Indiana: Indiana University Press. 1983.

————— 2008: *Maailman kielten erilaisuus ja samuus Osa 1*. Turku: Åbo akademis tryckeri. 2008.

Kielitoimiston sanakirja <https://www.kielitoimistonsanakirja.fi/netmot.exe?motportal=80> (6.2.2018).

Kommentoitu luettelo maahanmuuttajataustaisten opetuksen ja koulutuksen materiaaleista. Jokinen, P. – Immonen-Oikkonen, P. – Nissilä, L. (toim.). Opetushallitus. 2011.

KOPTJEVSKAJA-TAMM, MARIA & AHLGREN, INGER 2004: *NSM: Theoretical, methodological and applicational Problems*. Theoretical Linguistics, NSM-erikoisnumero 29. nro 3. 247–262.

KRASHEN, STEPHEN 1981: *Second language acquisition and second language learning*. Oxford: Pergamon.

LIGHTBOW, P. & WHITE, L. 1987: *The influence of linguistic theories on language acquisition research: description and explanation*. Language learning 37.: 483-510.

MCCARTHY, MICHAEL 1990: *Vocabulary*. Oxford: Oxford University Press.

MITCHELL, R & MYLES, F. 1998: *Second language learning theories*. Great Britain: Arnold. 1998.

MURRAY, D. W. & BUTTON, G. 1988: *Human emotions: Some problems of Wierzbicka's "simples"*. Blackwell Publishing.

NIDA, EUGENE 1975: *Componential analysis of meaning: An introduction to semantic structures*. The Hague: Mouton.

NIDA, EUGENE & TABER, CHARLES R. 1969/1982: *The theory and practice of translation*. Leiden: Brill. Pym, Anthony. 2010. *Exploring translation theories*. London: Routledge.

NSM:n kotisivu <https://intranet.secure.griffith.edu.au/schools-departments/natural-semantic-meta-language> (18.4.2018)

PAJUNEN, ANNELI 2018: Luentosarja SUOS1: *Suomi ja maailman kielet*. Tampereen yliopisto.

PEETERS, BRENT 2013: Language and Cultural Values – Towards an Applied Ethnolinguistics for the Foreign Language Classroom – Peeters, B., Mullan, K., Béal, C. (toim.) *Cross-culturally speaking, speaking cross-culturally*. Newcastle upon Tyne: Cambridge Scholars. 231-259.

PEIRCE, CHARLES S. 1932: *The Collected Papers of Charles Sanders Peirce, Vol. II: Elements of Logic*. Cambridge: Harvard University Press.: 230-231

PLATON: *Kratylos*. Teoksessa *Teokset II*. Toinen painos. Suomentanut Marja Itkonen-Kaila. Helsinki: Otava. 1999.

RIEMER, NICK 2006: *Reductive Paraphrase and Meaning: A Critique of Wierzbickian Semantics*. *Linguistics and Philosophy* 29.: 347-379

RUSSEL, BERTRAND 1946: *History of Western Philosophy*. Iso-Britannia: Unwin Brothers Limited Working.

SAPIR, EDWARD 1949: *Selected writings of Edward Sapir in language, culture and personality*. David Mandelbaum (toim.) Berkeley: University of California Press. 1949

SAUKKONEN, PAULI – HAIPUS, MARJATTA – NIEMIKOSKI, ANTERO – SULKALA, HELENA 1979: *Suomen kielen taajuussanasto*. Porvoo: WSOY. 1979.

Sisäministeriö: *Turvapaikanhakijat ja pakolaiset* <http://www.intermin.fi/fi/maahanmuutto/turvapaikanhakijat> (18.4.2018).

TANNER, JOHANNA 2012: *Rakenne tilanne ja kohteliaisuus. Pyynnöt S2-oppikirjoissa ja autenttissa keskusteluissa*. Tohtorinväitöskirja. Helsingin yliopisto. Suomen kielen, suomalais-ugrilaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos. Unigrafia Oy.

TUOVILA, SEIJA 2005: *Kun on tunteet: suomen kielen tunnesanojen semantiikkaa*. Tohtorinväitöskirja. Oulun yliopisto. Oulu University Press.

Uusi kieleemme – a blog about the Finnish language <https://uusikielemme.livejournal.com/5206.html> (1.4.2018).

VANHATALO 2014: Julkaisemattomat diat kurssilla ”SUOA6/LTLY210 Kulttuuriset skriptit ja kulttuuriset avainsanat. Näkökulmia pragmatiikkaan 2014-2015”. Tampereen yliopisto.

VANHATALO & TISSARI 2017: Esittelyssä alkusanakieli. *Virittäjä* 2/2017.: 244-262.

VANHATALO, ULLA – TISSARI, HELI – IDSTRÖM, ANNA —2014: Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish – Hamunen, M., Keisanen, T., Lantto, H., Lehti, L., Merke, S., (toim.) *SKY Journal of Linguistics*. The Linguistic Association of Finland. 67-94.

VEHKANEN, MARJUT 2015: *Kieliopista kommunikaatioon. Suomi toisena ja vieraana kielenä -oppikirjat vuosina 1866-1953*. Tohtorinväitöskirja. Helsingin yliopisto. Suomen kielen, suomalais-ugri-laisten ja pohjoismaisten kielten ja kirjallisuuksien laitos. Unigrafia Oy.

WHITE, LYDIA 1989: *Universal grammar and second language acquisition*. Amsterdam: John Benjamins. 1989

WHORF, BENJAMIN LEE 1956: *Language, thought, and reality: Selected writings*. John B. Carroll (toim.). New York: Wiley. 1956.

WIERZBICKA, ANNA 1972. *Semantic primitives*. Frankfurt: Athenäum.

————— 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins. 1988.

————— 1991: *Semantic complexity: Conceptual primitives and the principle of substitutability*. *Theoretical Linguistics* 17. 75-97.

————— 1992: *Semantics, culture and cognition: universal human concepts in culture-specific configurations*. Oxford: Oxford University Press.

————— 1997: *Understanding cultures through their key words: English, Russian, Polish, German, and Japanese*. Oxford: Oxford University Press. 1997.

————— 2010: *Experience, evidence, & sense: The hidden cultural legacy of English*. Oxford: Oxford University Press. 2010

————— 2014: *Imprisoned in English: The Hazards of English as a Default Language*. New York: Oxford University Press.

ZOBL, HELMUT 1983: *Markedness and the projection problem*. *Language learning* 33: 293-313.

9. LIITTEET

Liite 1) NSM:n suomenkielisten sanojen kaanon, eli alkusanat tyypillisissä käyttöyhteyksissään (Vanhatalo & Tissari 2017: *Virittäjän* verkkojulkaisu)

”

MINÄ

minä tahdon tehdä/tietää/sanoa jotakin, minä tahdon tätä, minä en tahdo tätä, minä en tiedä, jostain pahaa voi tapahtua minulle, joku kuten minä / minunlaiseni

SINÄ

minä tahdon sinun tekevän/tietävän/sanovan jotakin, jostain pahaa voi tapahtua sinulle, sinä olet minunlaiseni

JOKU ~ IHMINEN ~ HÄN

tämä ihminen, sama ihminen, joku toinen / toinen ihminen, tämä toinen ihminen

IHMISET

monet ihmiset, jotkut ihmiset, nämä ihmiset, ihmiset ajattelevat näin, ihmiset voivat sanoa – –, tietynlaiset ihmiset, monet ihmiset eivät voi tehdä jotain tällaista

JOKIN ~ ASIA

tämä asia, sama asia, jokin muu / jostain muuta, tämä toinen asia, jokin tietynlainen

RUUMIS ~ KEHO

jonkun ruumis/keho, ihmisten ruumiit/kehot, osa jonkun ruumiista/kehosta, eräänlainen ruumis/keho, kahdenlaisia ruumiita/kehoja, jostain pahaa tapahtuu jonkun ruumiissa/kehossa, joku tuntee jostain ruumiissaan/kehossaan

* * *

-LAINEN ~ -LÄINEN ~ LAJI

tällainen, samanlainen, toisenlainen, tämä toisenlainen, jokin/joku eräänlainen, eräänlaiset/kahdenlaiset/monenlaiset ihmiset

OSA

jonkun ruumiin osa, tämä osa, sama osa, toinen osa, tämä toinen osa, tämänlaisella asialla / tämänlaisessa asiassa on kaksi/monta osaa

* * *

TÄMÄ ~ SE

tämä joku (jokin), nämä ihmiset, tällainen, tämä osa, tänä aikana, tässä paikassa, tämän takia, se on kuten tämä / näin: – –

SAMA

sama ihminen, sama asia, sama osa, samanlainen, samaan aikaan, samassa paikassa, joku sanoo/tekee/ajattelee/tietää/tahtoo/tuntee samaa (samoin, saman)

TOINEN ~ MUU

joku muu/toinen, jokin muu, muuna/toisena aikana, jossain muualla/toisaalla, muut osat, muunlaiset/toisenlaiset, tämä toinen osa, tämä toisenlainen, tämä toinen ihminen, tämä toinen asia

* * *

YKSI ~ ERÄS

yksi/eräs ihminen, yksi/eräs asia, yksi/eräs osa, yhdenlainen/eräänlainen, yhdessä/eräässä paikassa, yhteen aikaan / eräänä aikana, yksi/eräs näistä asioista/ihmisistä, jotain yhdenlaista/eräänlaista, yksi/eräs asia lisää

KAKSI

kaksi asiaa, kaksi osaa, kahdenlaiset/kahdenlaisia, kaksi näistä asioista/ihmisistä, kaksi asiaa lisää

JOKIN ~ JOKU ~ MUUTAMA

jotkut/muutamat ihmiset, jotkin/muutamat asiat, jotkin/muutamat osat, jonkinlaiset/muutamanlaiset, joinakin aikoina, joissakin/muutamissa paikoissa jotkut/muutamat näistä asioista / ihmisistä, jotain tästä tällaisesta (esim. vedestä)

KAIKKI

kaikki ihmiset, kaikki asiat, kaikki osat, kaikenlaiset, kaikkina aikoina, kaikissa paikoissa, kaikki tästä tällaisesta (esim. vedestä)

PALJON ~ MONI

paljon ihmisiä / monet ihmiset, paljon asioita / monet asiat, paljon osia / monet osat, monenlaisia, monina aikoina, monissa paikoissa, paljon jotain tällaista (esim. vettä), paljon enemmän

VÄHÄN ~ HARVA

vähän ihmisiä / harvat ihmiset, vähän asioita, vähän jotain tällaista (esim. vettä)

* * *

HYVÄ

jotain hyvää, joku hyvä, jotain hyvää tapahtuu, tehdä jotain hyvää (jollekulle), tuntea jotain hyvää, tämä on hyvä(ä), on hyvä jos —

PAHA

jotain paha, joku paha, jotain paha tapahtuu, tehdä jotain paha (jollekulle), tuntea jotain paha, tämä on paha(a), on paha jos —

* * *

ISO ~ SUURI

jokin suuri, iso paikka, iso/suuri osa

PIENI

jokin pieni, pieni paikka, pieni osa

* * *

AJATELLA

joku ajattelee jotakuta/jotain, joku ajattelee hyvää/pahaa jostakusta/jostakin, joku ajattelee

näin: ”– –”, monet ihmiset ajattelevat näin: ”– –”

TIETÄÄ

joku tietää jotakin (monia asioita) jostakusta/jostakin, ihmiset voivat tietää tämän, tiedän että – –

TAHTOA

tahdon tätä, joku tahtoo jotakin, joku tahtoo tehdä/tietää/sanoa jotakin, joku tahtoo

jonkun muun tekävän/tietävän/sanovan jotakin, joku tahtoo jotakin tapahtuvan, tahdon sitä hyvin paljon

EI TAHTOA

en tahdo tätä, joku ei tahdo tätä, joku ei tahdo tehdä jotain, joku ei tahdo jotain tapahtuvan

TUNTUA

jostakin tuntuu joltakin (hyvältä/pahalta), joku tuntee jotakin (hyvää/pahaa) osassa ruumistaan, jostakin tun-
tuu tältä, jostakin tuntuu hyvältä/pahalta jotakuta toista kohtaan

NÄHDÄ

joku näkee jonkun/jotakin (jossain paikassa), ihmiset voivat / eivät voi nähdä hyvin

tässä paikassa, joku/i ihmiset voi(va)t/ei(vät) voi nähdä tätä

KUULLA

joku kuulee jotain, ihmisistä voi tuntua pahalta kun he kuulevat tämän sanan

SANOA

minä sanon: – –, joku sanoo jotain (hyvää/pahaa) jollekin, joku sanoo jotain (hyvää/

pahaa) jostakusta/jostakin, joku sanoo jotain tällaista: ”– –”, joku sanoo jotain sanoilla, joku sanoo sanan jol-
lekin

SANAT

monet sanat / paljon sanoja, toiset sanat, yksi sana, tietynlaiset sanat, sanoa jotakin sanoilla/sanoin/sanoitta,
sanoa nämä sanat, nämä sanat sanovat jotakin

TOSI ~ TOTTA

tämä on totta, tämä ei ole totta

* * *

TEHDÄ

joku tekee jotakin (jollekulle muulle), joku tekee jotakin jollakin / osalla kehoaan/ruumistaan, joku tekee jo-
takin jonkun muun kanssa, joku tekee jotakin hyvää/pahaa (jollekulle muulle)

TAPAHTUA

jotain tapahtuu, jotain tapahtuu jollekulle, jotain tapahtuu jollekin, jotain tapahtuu

jossain (paikassa)

LIKKUA

joku liikkuu (tässä paikassa), jokin liikkuu tässä paikassa, tämän ihmisen ruumiin osat liikkuvat kuten tämä
ihminen tahtoo

KOSKEA

jokin koskee jotain muuta (jossain), jokin koskee osaa jonkun ruumiissa/kehossa, joku/i ihmiset voi(vat) / ei(vät) voi koskea jotakin

* * *

OLLA ('olla jossakin')

joku on jossain (paikassa), jokin on jossain (paikassa), joku on jonkun muun kanssa

OLLA ('olla olemassa')

tässä paikassa on jotain, tässä paikassa on joku, on kahdenlaisia/monenlaisia – –

OLLA ('olla joku/jokin')

tämä ihminen on minunlaiseni, tämä on jokin (jotain) tietynlainen/eräänlainen (tietynlaista/eräänlaista), tämä on jokin suuri/pieni (suurta/pientä), tiedän kuka tämä ihminen on

OLLA ('olla jonkun')

tämä asia (veitsi, paita, jne.) on minun, tämä asia on jonkun muun, kenen (veitsi, paita, jne.) tämä on?

* * *

ELÄÄ

joku elää kauan aikaa, joku elää tässä paikassa, monet ihmiset elävät tässä paikassa,

tämä ihminen elää jonkun toisen ihmisen kanssa, on hyvä jos joku elää näin

KUOLLA

joku kuolee tänä aikana, kaikki ihmiset kuolevat joskus

* * *

MILLOIN ~ JOLLOIN ~ SILLOIN ~ AIKA ~ KUN

tämä aika / tänä aikana, sama aika / samaan aikaan, toisena/muuna aikana, sinä toisena aikana, joinakin aikoina, monina aikoina, sinä aikana kun – –

NYT

nyt täällä tapahtuu jotain, kun nyt sanon jotain

ENNEN ~ AIKAA SITTEN ~ AIKAISEMMIN

ennen tätä, jonkin aikaa sitten, vähän aikaa sitten / aikaisemmin, kauan aikaa sitten / aikaisemmin

JÄLKEEN ~ AJAN KULUTTUA

tämän jälkeen, jonkin ajan kuluttua, vähän ajan kuluttua, pitkän ajan kuluttua

KAUAN (AIKAA) ~ PITKÄN AJAN

kauan (aikaa) sitten, pitkän ajan kuluttua, hyvin pitkä aika, hyvin kauan (= jonkin aikaa, pitkän aikaa)

VÄHÄN AIKAA ~ LYHYEN AJAN

vähän aikaa sitten, vähän ajan kuluttua, hyvin lyhyt aika, hyvin vähän aikaa

(= jonkin aikaa, vähän aikaa)

JONKIN AIKAA

jokin aika sitten, jonkin ajan kuluttua, sitä tapahtuu jonkin aikaa, joku tekee tätä jonkin aikaa, tämän ajan kuluessa (= tänä aikana, jonkin aikaa)

HETKI

se tapahtuu hetkessä, tänä hetkenä

* * *

MISSÄ ~ JOSSA ~ PAIKKA ~ -SSA

tämä paikka / tässä paikassa, sama paikka / samassa paikassa, jossain muualla, tämä
toinen paikka / tässä toisessa paikassa, joissain paikoissa, monissa paikoissa, paikassa
jossa — —

TÄSSÄ ~ TÄÄLLÄ

täällä tapahtuu nyt jotain

PÄÄLLÄ ~ YLÄPUOLELLA

tämän paikan päällä/yläpuolella, kaukana tämän paikan yläpuolella, joku muiden ihmisten yläpuolella

ALLA ~ ALAPUOLELLA

tämän paikan alla/alapuolella, kaukana tämän paikan alapuolella

KAUKANA

kaukana tästä paikasta

LÄHELLÄ

tämän paikan lähellä, jonkun lähellä

PUOLI

tällä puolella, samalla puolella, yhdellä puolella, kahdella puolella, kaikilla puolilla

SISÄ

tämän jonkin sisällä, ihmisen sisällä, jonkun ruumiin/kehon sisäosa

* * *

EI ~ OLLA ...MATTA

ei hyvä, ei paha, ei näin/tällainen, minä en tiedä, joku ei voi tehdä tätä

EHKÄ

se on ehkä näin, se ei ole ehkä näin, ehkä joku muu voi tehdä sen

VOIDA

joku voi (ei voi) tehdä jotain, joku ei voi olla tekemättä jotain, jotain (hyvää/pahaa) voi tapahtua, se voi olla
näin: — —

KOSKA ~ VUOKSI ~ TAKIA

tämän vuoksi/takia, se tapahtui koska tämä ihminen teki jotain aikaisemmin, — — ei
minkään muun vuoksi

JOS

jos tätä tapahtuu jonkin aikaa, — —, jos sinä teet tämän, — —, jos joku tekee jotain tällaista,

— —

* * *

HYVIN ~ ERITTÄIN

hyvin/erittäin suuri, hyvin/erittäin pieni, erittäin hyvä, hyvin/erittäin paha, hyvin/
erittäin kaukana, hyvin/erittäin lähellä, hyvin/erittäin vähän aikaa, hyvin/erittäin
kauan aikaa

ENEMMÄN ~ LISÄÄ (~ ENÄÄ)

joku tahtoo enemmän/lisää, joku tekee enemmän, joku tahtoo tietää/sanoa/ajatella
enemmän siitä, yksi lisää/enemmän, kaksi lisää/enemmän, monta lisää/enemmän, ei
elä enää, ei enää tällainen

* * *

NÄIN ~ KUTEN ~ TÄLLÄ TAVALLA

tapahtuu näin: —, se on kuten tämä: —, joku ajattelee tällä tavalla: ”—”, — tekee
sen näin / tällä tavalla, ei toisin, joku tekee sen näin, tapahtuu kuten tämä ihminen tahtoo, joku kuten minä /
minunlaiseni

”